Etablierung des stickstofffixierenden Alphaproteobakteriums *Sinorhizobium meliloti* als Chassis-Organismus in der Synthetischen Mikrobiologie

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

dem Fachbereich Biologie

der Philipps-Universität Marburg

vorgelegt im November 2016

von Johannes Döhlemann

geboren in Adenau

Die Untersuchungen zur vorliegenden Arbeit wurden von April 2013 bis November 2016 unter der Betreuung von Frau Prof. Dr. Anke Becker am LOEWE-Zentrum für Synthetische Mikrobiologie in der Arbeitsgruppe Vergleichende Genomik in Marburg durchgeführt.

Vom Fachbereich Biologie

der Philipps-Universität Marburg als Dissertation

angenommen am:

Erstgutachter: Frau Prof. Dr. Anke Becker

Zweitgutachter: Herr Prof. Dr. Martin Thanbichler

Tag der mündlichen Prüfung:

Aus dieser Dissertation hervorgegangene Publikationen:

1. Frage, B; Döhlemann, J; Robledo, M; Lucena, D; Sobetzko, P; Graumann, PL; Becker, A (2016). Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. *Mol Microbiol.*100(5):808-23

2. Döhlemann, J; Brennecke, M; Becker; A (2016). Cloning-free genome engineering in *Sinorhizobium meliloti* advances applications of Cre/*loxP* site-specific recombination. *J Biotechnol.* 233:160-70

3. Döhlemann, J*; Wagner, M*; Happel, C; Sobetzko, P; Becker A. A family of single copy *repABC*-type shuttle vectors stably maintained in the alpha-proteobacterium *Sinorhizobium meliloti*. *In Revision*
   *contributed equally

Abkürzungsverzeichnis

Abb.	Abbildung
AS	Aminosäure
bp	Basenpaare
BGM	*Bacillus subtilis genome*
BreT	*Bacillus recombinatorial transfer*
ca.	circa
CFGE	*cloning-free genome editing*
Chr	Chromosom
CRISPR	*clustered regularly interspaced short palindromic repeats*
crRNA	CRISPR-RNA
DNA	Desoxyribonukleinsäure
DR	*directed repeats*
DSB	DNA-Doppelstrangbruch
dsDNA	doppelsträngige DNA
ESB	DNA-Einzelstrangbruch
et al.	*et alii*
FROS	*Fluorescent repressor operator system*
gen.	genomisch
HDR	*homology directed repair*
Hi-C	*Chromosome-Conformation-Capture*
HR	homologe Rekombination
kb	Kilobasenpaar
LE	linkes Element von *loxP*
lR	linkes Replichor
MAGE	*multiplex automated genome engineering*
Mb	Megabasenpaar
mRNA	Matrizen-RNA
N_2	atmosphärischer Stickstoff
a/cNHEJ	*alternative/classical non-homologous End Joining*
nt	Nukleotid
Nod	Nodulation
oriV	Replikationsursprung
QS	*quorum sensing*
RBS	Ribosomenbindestelle
RE	rechtes Element von *loxP*
RM-System	Resitriktions-Modifikationssystem
RMCE	*recombinase mediated cassette exchange*
RNA	Ribonukleinsäure
rR	rechtes Replichor
ssDNA	einzelsträngige DNA
SSB-Protein	ssDNA-bindendes Protein
sp.	Spezies
TALEN	*transcription activator-like effector nuclease*
TAR	*transformation-associated recombination*
terA/B/C	Terminus-Region von pSymA/ pSymB/ des Chromosoms
Tab.	Tabelle
TSS	Transkriptionsstart
UTR	untranslatierte Region
Wt	Wildtyp
z.B.	zum Beispiel
ZFN	Zinkfinger-Nukleasen

Inhaltsverzeichnis

1. Zusammenfassung

Die Synthetische Biologie benötigt Strategien, die effizientes *Genetic Engineering* erlauben. Eine grundlegende Voraussetzung hierzu ist die Etablierung von Techniken zur spezifischen DNA-Modifikation im Umfang einzelner Nukleotide bis hin zu ganzen Chromosomen. Während für Modellorganismen der Synthetischen Mikrobiologie ein großes Repertoire leistungsfähiger Werkzeuge zur Verfügung steht, war das Methodenspektrum zur genetischen Veränderung von Rhizobien aus der Klasse der Alphaproteobakterien lange Zeit eingeschränkt.

In symbiotischen Lebensgemeinschaften mit Leguminosen besitzt *Sinorhizobium meliloti* die Fähigkeit, atmosphärischen Stickstoff zu biogenen Stickstoffverbindungen zu assimilieren, wodurch dem α-Rhizobium insbesondere landwirtschaftliche Relevanz zukommt. Aufgrund seiner besonderen Genomarchitektur, die neben einem Chromosom ein weiteres Chromid (pSymB) und Megaplasmid (pSymA) aufweist, ist *S. meliloti* außerdem ein Modellorganismus zur Erforschung mehrteiliger Genome in Bakterien. pSymA und pSymB, die ~45% der genomischen DNA tragen, besitzen *repABC*-basierte Replikationsursprünge, welche die Integration dieser extra-chromosomalen Replikons in den Zellzyklus vermitteln.

Die *repABC*-Region wurde in dieser Arbeit als funktionelle Einheit identifiziert, die Chromosomen-ähnliche Eigenschaften wie eine stabile Vererbung und einfache Kopienzahl vermittelt. Auf Basis heterologer *repABC*-Kassetten verschiedener α-Rhizobien konnte außerdem ein *Shuttle*-Vektorsystem etabliert werden. Ein modulares und standardisiertes Assemblierungssystem ermöglichte hierbei die *in-vitro* Herstellung artifizieller Mini-Replikons. Diese können sowohl als klassische Klonierungsvektoren in *E. coli*, aber auch zur Etablierung mehrteiliger Expressionssysteme in *S. meliloti* eingesetzt werden. Eine neue *Genome Editing*-Methode sowie verbesserte Cre/*lox*-basierte Rekombinationsstrategien wurden entwickelt und erweitern das Repertoire genetischer Werkzeuge für *S. meliloti*. Ihre Anwendung konnte mittels Deletion und Inversion umfangreicher Genomabschnitte, sowie substantieller Genom-Umstrukturierungen demonstriert werden. Cre/*lox*-Anwendungen und eine Methode zur *in-vivo* Klonierung großer DNA-Fragmente ermöglichen neue experimentelle Ansätze, die bei der mechanistischen Aufklärung des mehrteiligen Genoms sowie in synthetischen Anwendungen von großem Nutzen sind. Die robuste Funktion des Cre/*lox*-Systems und die weite Verbreitung *repABC*-basierter Replikons innerhalb der α-Rhizobien bergen zudem das Potential, die in dieser Arbeit etablierten Methoden auf weitere Vertreter dieser landwirschaftlich und biotechnologisch relevanten Organismen zu übertragen.

Summary

Synthetic biology requires strategies, which facilitate efficient genetic engineering. DNA editing in the range of single nucleotides up to whole genomes requires appropriate molecular methods. While powerful tools are available for prokaryotic model organisms, the repertoire for genetic engineering of rhizobia of the class alphaproteobacteria is still limited.

The α-rhizobium *Sinorhizobium meliloti* is able to fix atmospheric nitrogen in symbiotic relationships with legumes. Therefore it is of high agricultural value. Due to its remarkable genome architecture, which besides a main chromosome exhibits a chromid (pSymB) and a megaplasmid (pSymA), *S. meliloti* emerged as model organism for studying multipartite genomes in bacteria. pSymA and pSymB carry ~45% of the genomic content and are based on *repABC*-type replication origins, which integrate these extra-chromosomal replicons into the cell cycle. The *repABC* region of pSymA was further identified to confer the chromosome-like features as single copy replication and stable inheritance to artificial DNA constructs.

In this work, heterologous *repABC* cassettes derived from different rhizobial species were used to establish a novel shuttle-vector system for *S. meliloti*. This highly modular and standardized assembly system facilitates *in-vitro* construction of artificial mini replicons. Resulting pABC vectors serve as *E. coli*-compatible cloning vehicels and enable establishment of complex, single-copy expression systems. Furthermore, a cloning-free genome editing method and an optimized Cre/*lox* recombination system were developed and expand the genetic toolbox of *S. meliloti*. Respective engineering approaches were demonstrated by substantial genome rearrangements and large scale deletion and inversion systems. Cre/*lox* recombination strategies and an inducible *repABC*-type replication origin further facilitated *in-vivo* cloning, which enables new experimental approaches both in genome biology, as well as for synthetic applications. The robust function of Cre/*lox* and the wide distribution of *repABC*-based replicons among the alphaproteobacteria likely allow an application of the established methods in additional organisms to exploit their genetic resources for agricultural and biotechnological purposes.

2. Einleitung

2.1 Synthetische Biologie

Seit der ersten vollständigen Genomsequenzierung des Bakteriums *Haemophilus influenza* im Jahre 1995 erfuhren Techniken zur DNA-Synthese, -*in-vitro*-Assemblierung und -Sequenzierung eine rapide Entwicklung (Fleischmann *et al*., 1995; Kosuri und Church, 2014; Gibson *et al*., 2009; Kok *et al*., 2014; Werner *et al*., 2012; van Dijk *et al*., 2014; Ke *et al*., 2016). Leistungsstarke analytische Methoden bereiteten zudem das Feld der Genomik, Transkriptomik, Proteomik und Metabolomik, stets begleitet von computerunterstützten Plattformen, die die Verarbeitung der anfallenden Datenmengen ermöglichten. Diese technischen Fortschritte wirkten als Katalysator bei der Entstehung einer neuen biologischen Disziplin, die weit über das Wirkungsfeld der klassischen Biotechnologie hinaus geht. Die Synthetische Biologie speist sich aus der Grundlagenforschung etablierter biologischer und biochemischer Domänen, bedient sich bioinformatischer und physikalisch-mathematischer Methoden, und kombiniert nach ingenieurstechnischen Konstruktionsprinzipien sowohl synthetische, als auch biologische Komponenten. Das Ziel ist die Konstruktion neuer biologischer Systeme mit anwendungsnahem Charakter. Beispiele erfolgreicher Anwendungen finden sich nicht nur in der Mikrobiologie (Wu *et al*., 2016). Sie reichen von pflanzlichen Modellorganismen, in welchen synthetische Sensorsysteme zur Detektion von bakteriellen Pathogenen, Elicitoren oder Nematoden entwickelt wurden (Nemhauser und Torii, 2016), bis hin zu synthetischen Netzwerken in menschlichen Zellen, in welchen sie zur Therapie metabolischer und immunologischer Krankheiten beitragen (Kis *et al*., 2015). Der Fokus dieser Arbeit soll jedoch auf dem Bereich der Synthetischen Mikrobiologie liegen.

Pionierarbeiten auf dem Gebiet der Synthetischen Biologie gehen auf die Entwicklung genetischer Schaltkreise in Prokaryoten zurück (Elowitz und Leibler, 2000; Gardner *et al*., 2000; Fung *et al*., 2005), und theoretische Modelle wurden entworfen, die die Regulation und Dynamik dieser Systeme beschreiben (He *et al*., 2016) und in logische Operatoren der Boolschen Algebra übersetzen (Anderson *et al*., 2007; Friedland *et al*., 2009; Wang *et al*., 2011a). Die Entwicklung von Strategien zur Kontrolle des zellulären Metabolismus ist integraler Bestandteil der Synthetischen Mikrobiologie mit positiven Implikationen auf biotechnologische, pharmazeutische und biomedizinische Anwendungen (Padilla-Vaca *et al*., 2015; Braff *et al*., 2016). *Metabolic Engineering* ist ein Schlüsselbegriff, der die Optimierung nativer sowie die Etablierung heterologer Biosythesewege in geeigneten Organismen beschreibt. Bedeutende Chassis-Organismen sind *Escherichia coli*, *Saccharomyces cerevisiae*, und Vertreter der Gattungen *Streptomyces*, *Bacillus* und *Corynebacterium* (Wu *et al*., 2016). Sie ermöglichen beispielsweise die *de-novo* Synthese einer Vielzahl von Flavonoiden und dienen somit der Herstellung medizinisch wirksamer Lebensmittel und Kosmetika (Pandey *et al*., 2016). Synthetisch-genetische

Netzwerke unterstützen die Arzneimittelforschung (Kis *et al*., 2015) und ermöglichen das Design synthetischer Bakteriophagen zur Behandlung von Infektionskrankheiten, einem alternativen Ansatz zur aufwändigen Entwicklung neuer Antibiotika (Citorik *et al*., 2014). Auch die biosynthetische Herstellung von Biotreibstoffen ist ein zentraler Aspekt des *Reverse Engineering* aufwendiger Stoffwechselwege. So eignen sich Spezies aus der Klasse *Clostridia* aufgrund ihres großen Inventars an Hydrolasen besonders zur Degradierung von Biomasse, um beispielsweise Cellulose zu Ethanol zu konvertieren (Bayer *et al*., 2008; Lynd *et al*., 2008; Dam *et al*., 2011). Zuletzt wurde *Pseudomonas putida* als physikalisch und metabolisch robuster Chassis-Organismus etabliert und erfolgreich zur Produktion von Biotreibstoff, verschiedener Feinchemikalien und phenolischer Polymere eingesetzt (Nikel *et al*., 2016). Die Weiterentwicklung des *Pathway Engineering* mündet schließlich im Design maßgeschneiderter Produktionsstämme und strebt den Aufbau komplexer, biosynthetischer Stoffwechselwege an (Ng *et al*., 2015). Eine Vielzahl bioinformatischer Plattformen wurde entwickelt, um auf Basis bekannter Biosyntheseprozesse und umfangreicher Reaktionsdatenbanken optimierte Systeme zu modellieren, die beispielsweise eine synthetische CO_2-Fixierung ermöglichen (Schwander *et al*., 2016; Bordbar *et al*., 2014; Bar-Even *et al*., 2010).
Die Notwendigkeit eines umfassenden Verständnis der zugrundeliegenden Genetik sowie die Verfügbarkeit leistungsfähiger, genetischer Werkzeuge und adäquater DNA-Vehikel ist offenkundig. *Genetic Engineering* ist die Grundlage der Synthetischen Biologie.

2.1.1 Genetische Werkzeuge zur *in-vivo* Modifikation einzelner Gene und bakterieller Genome

Die Konstruktion synthetischer Schaltkreise und Biosynthesewege, aber auch die Entwicklung optimierter Chassis-Organismen mit reduziertem oder Minimalgenom bedarf leistungsfähiger Werkzeuge, mittels derer DNA-Moleküle *in-vivo* auf spezifische Weise modifiziert werden können (*Genome Editing*). Im Folgenden wird ein Überblick über die Funktionsweise verfügbarer Werkzeuge und Methoden gegeben, die eine Bearbeitung der DNA im Maßstab einzelner Nukleotide bis ganzer Chromosome ermöglichen.

2.1.1.1 Methoden des ortsspezifischen *Genome Editings*

Die am weitesten verbreitete Methode zur genetischen Modifikation basiert auf der genomischen Integration heterologer DNA. Der Mechanismus, der sich hierbei zu Nutze gemacht wird, beruht auf einer endogenen DNA-Reparatur-Maschinerie, die DNA-Doppelstrangbrüche (DSBs) und blockierte Replikationsgabeln mittels homologer Rekombination (HR) wiederherstellt (Bell und Kowalczykowski, 2016; Court *et al*., 2002; Yeeles und Dillingham, 2010). Solche HR-Ereignisse, die am detailliertesten an *E. coli* studiert wurden, sind in der Regel RecA-abhängig und werden durch das RecBCD-System

initiiert. RecBCD ist ein Heterotrimer mit DNA-Helikase-, ssDNA-Endonuklease- und ss/dsDNA-Exonuklease-Aktivität. Der RecBCD-Rekombinationskomplex erkennt DSBs und prozessiert dsDNA so lange, bis er auf eine 8 bp große *chi*-Sequenz in korrekter Orientierung trifft. Dieser Prozess führt im Falle von Fremd-DNA, beispielsweise Bakteriophagen-DNA, zur DNA-Degradierung und kann zudem die Adaptation mancher *clustered regularly-interspaced short palindromic repeat* (CRISPR)-Systeme unterstützen (Levy *et al*., 2015). Im Zuge der HR wird der prozessierte 3'-DNA-Einzelstrang schnell von ssDNA-bindenden (SSB-) Proteinen bedeckt. RecA wird daraufhin mithilfe des RecFOR- oder RecOR-Komplexes auf die ssDNA geladen, wodurch die Bildung des SSB-ssDNA Nukleoproteins gestört und die ATP-abhängige Polymerisierung von RecA eingeleitet wird (Bell und Kowalczykowski, 2016). Das präsynaptische Filament ist das RecA-Assemblierungsprodukt, welches einerseits eine SOS-Antwort und infolgedessen eine Verzögerung des Zellzyklus induziert, und andererseits einen 8-Nukleotid-Suchprozess zur Identifizierung einer homologen DNA-Sequenz initiiert. So genügen in *E. coli* und *S. cerevisiae* 12 bzw. 13 Nukleotide, um eine homologe Sequenz zu definieren und die HR einzuleiten. Fälschlicherweise wurde der RecBCD-Pfad in Gram-negativen Bakterien lange Zeit als obligatorisch betrachtet, während das funktionale Homolog AddAB exklusiv Gram-positiven Bakterien wie *Bacillus subtilis* zugeschrieben wurde (Chedin und Kowalczykowski, 2002). Tatsächlich konnten AddAB-Derivate mittlerweile in zahlreichen Vertretern der Alphaproteobakterien wie *Caulobacter crescentus* und *Sinorhizobium meliloti* identifiziert werden (Zuñiga-Castillo *et al*., 2004).

Während die Häufigkeit homologer Rekombinationsereignisse in Eukaryoten mit einer Frequenz von 10^{-6} sehr gering ist und HR-basierte Anwendungen zur genetischen Modifizierung daher wenig praktikabel sind (Nafissi und Slavcev, 2014), stellt die HR in der Synthetischen Mikrobiologie den Mechanismus der Wahl dar, um genomische Modifikationen bzw. die Integration heterologer DNA zu bewerkstelligen (Court *et al*., 2002). Zur Optimierung der Integrationsfrequenz kommen außerdem oftmals RecA-unabhängige Alternativen zum Einsatz, die auf Bakteriophagen-stämmigen Rekombinationssystemen beruhen. So findet das auf dem Bakteriophagen Lambda beruhende Lambda-Red-System insbesondere in *Genetic Engineering*-Prozeduren in *E. coli* (Poteete *et al*., 2001; Lee *et al*., 2009), aber auch in *Salmonella* und *Streptomyces* Verwendung (Gust *et al*., 2004; Karlinsey *et al*., 2007). Auch die enorme Effizienz des nativen Rekombinationssystems von *S. cerevisiae* und sein Potential in biotechnologischen Anwendungen wurde früh erkannt (Oldenburg *et al*., 1997) und bildet heute die Grundlage der leistungsfähigsten Assemblierungssysteme zur Fusion multipler DNA-Fragmente (siehe Kap. 2.1.2).

Nicht-homologes *End Joining* (NHEJ) ist ein zur HR komplementärer DNA-Reparaturmechanismus, der auf einer Ligase-vermittelten Fusion freier DNA-Enden beruht und bereits 1982 in Säugetierzellen entdeckt wurde (Wilson *et al*., 1982). Während alternatives NHEJ (aNHEJ) bislang auf Eukaryoten beschränkt zu sein scheint, wurden Komponenten des klassischen NHEJ (cNHEJ) mittlerweile auch in

Prokaryoten der Gattung *Bacillus*, *Mycobacterium* und *Sinorhizobium* identifiziert (Deriano und Roth, 2013; Kobayashi *et al*., 2008).

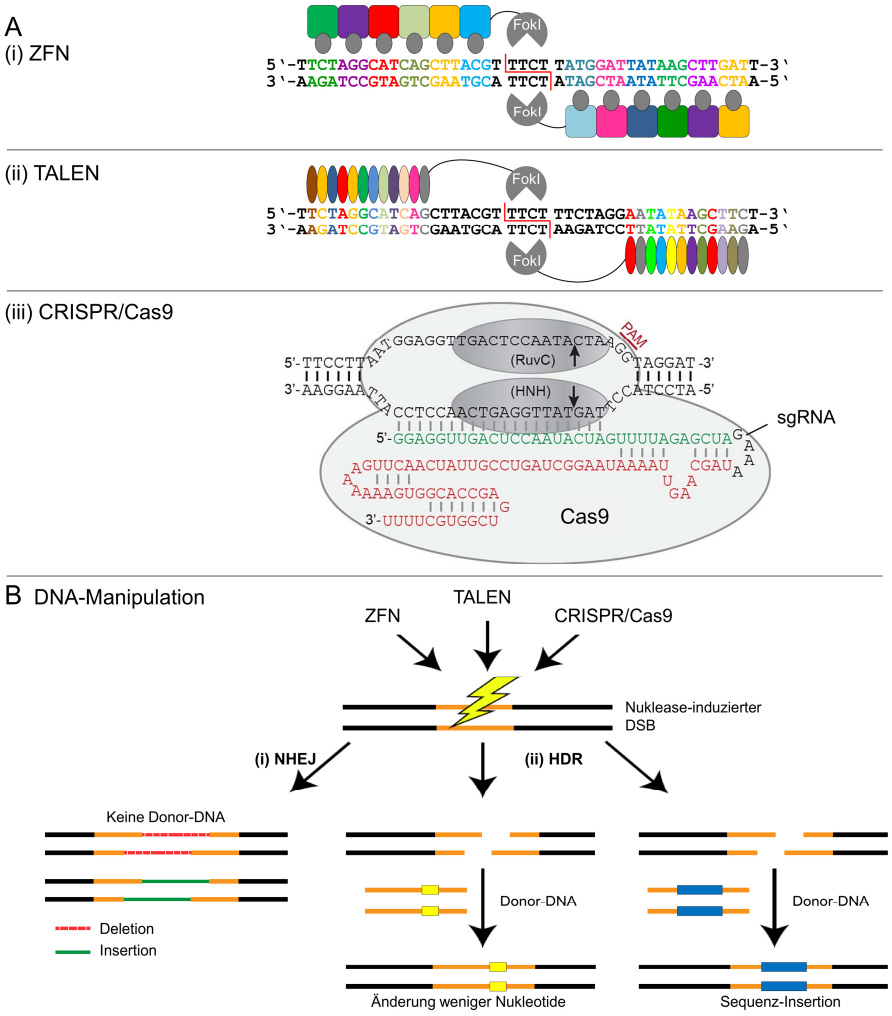


Abbildung 2.1: Grundlegende *Genome Editing*-Ansätze zur lokalen Veränderung genomischer DNA. **A:** Werkzeuge zur Einführung von ortsspezifischen DSBs. (i) Zinkfingernukleasen (ZFN) binden spezifisch drei Nukleotide. Fusioniert an zwei ZF-Arrays, die denselben Locus attackieren, verursacht die dimerisierte Nuklease FokI einen DSB. (II) Auch *transcription activator-like* Effektoren (TALE) können mit FokI fusioniert werden (TALEN) und als individuelle Effektorproteine die Einführung spezifischer DSBs vermitteln. (iii) Die Nuklease Cas9 wird in einem modifizierten CRISPR-System mittels sgRNA, die sich aus crRNA (grün) und tracrRNA (rot) zusammensetzt, zu einer Ziel-Sequenz geleitet. Die Ziel-DNA hybridisiert dabei mit der ~20 bp großen *Protospacer*-Sequenz. Sofern eine *protospacer adjacent motif* (PAM)-Sequenz (NGG) unmittelbar stromabwärts der Ziel-DNA vorliegt, schneiden die Nuklease-Domänen RuvC und HNH Leit- und Folgestrang 4 bp stromaufwärts der PAM-Sequenz (Pfeile). **B:** Induzierte DSBs ermöglichen ortsspezifisches *Genome Editing* mittels NHEJ oder *homology-directed repair* (HDR)-basierter Reparaturmechanismen. (i) Die NHEJ-vermittelte Reparatur des betroffenen Lokus führt zu kleinen Deletionen oder Insertionen (Mojica und Montoliu, 2016) und wird primär in Eukaryoten (Sonoda *et al*., 2006) zur Gen-Inaktivierung genutzt. (ii) Durch die simultane Einführung eines DSBs und exogener DNA, die zu den entstehenden DNA-Enden homolog ist (orange), kann eine HR-vermittelte Mutation (gelb), bzw. Insertion größerer DNA-Sequenzen (blau) erzielt werden. Der priorisierte Modus ist Organismus-spezifisch. Modizifizert nach MacDonald & Deans (2016) und Waaijers & Boxem (2014).

HR- und NHEJ-basierte DNA-Reparaturmechanismen sind die Grundlage des *Genome Editings* (Abb. 2.1). Zur Steigerung der Effizienz genetischer Modifikationen werden zudem DSB- oder DNA-Einzelstangbruch (ESB)-vermittelnde Werkzeuge eingesetzt, die die Integrationsfrequenz simultan eingebrachter Donor-DNA steigern.

Die Möglichkeit, Transkriptionsfaktoren mit Zinkfinger (ZF)-Domäne zu artifiziellen Restriktionsenzymen umzufunktionieren, wurde bereits in den frühen 1990er Jahre erkannt (Desjarlais und Berg, 1992, 1993). Cys_2His_2-ZF gehören zu den wichtigsten Vertretern dieser DNA-Bindeproteine, die mittels variabler Aminosäure (AS)-Reste drei Nukleotide umfassende Motive erkennen und spezifisch binden (Hossain *et al*., 2015). Oft werden sechs ZF-Derivate zur Bindung einer 18 bp großen Ziel-DNA fusioniert und mit der Endonuklease FokI verknüpft (Abb. 2.1A i). Resultierende ZF-Nukleasen (ZFN) vermitteln die Einführung ortsspezifischer DSBs, deren anschließende HR- oder NHEJ-basierte Reparatur zur Einführung spezifischer Deletionen und Insertionen genutzt wird.

Obwohl mit der Ros/MucR-Familie mittlerweile ZF-ähnliche Transkriptionsfaktoren prokaryotischen Ursprungs entdeckt wurden (Malgieri *et al*., 2015), basieren typische ZFN auf DNA-Bindeproteinen eukaryotischer Organismen und kommen auch vornehmlich in diesen zum Einsatz. *Off-target* Effekte, das aufwändige Design synthetischer ZFNs (Hossain *et al*., 2015), sowie die Entdeckung leistungsfähigerer *Editing*-Systeme lassen diese Methode zudem zusehends in den Hintergrund rücken.

Der *transcription activator-like* Effektor (TALE) ähnelt in technischer Hinsicht den ZF-Proteinen, da sich auch seine DNA-Bindungseigenschaften mittels variabler AS-Reste umprogrammieren lassen (Boch *et al*., 2009; Moscou und Bogdanove, 2009). Der aus dem Pflanzenpathogen *Xanthomonas* stammende Transkriptionsregulator wurde im Jahr 2009 dechiffriert und umgehend als Werkzeug für synthetisch-genetische Anwendungen etabliert (MacDonald und Deans, 2016). Eine TALE DNA-Bindedomäne besteht aus 7-34 Wiederholungssequenzen (*directed repeats*, DR), die sich aus jeweils 33-35 AS zusammensetzen. AS-Reste in Position 12 und 13 kodieren dabei die Bindespezifität des DR, und der modulare Aufbau erlaubt die Synthese artifizieller, ~500-700 AS großer Effektoren zur Bindung beliebiger DNA-Sequenzen. Auch TALE-basierte *Genome Editing*-Methoden beruhen auf der Fusion mit einer DNA-Nuklease (TALEN), die ortsspezifische DSBs vermittelt (Abb. 2.1A ii). Die Herstellung artifizieller TALENs ist jedoch aufwändig und mangels technischer Alternativen primär in der Eukaryotengenetik von Relevanz.

Die Entdeckung des CRISPR-Cas Systems und seine Anwendung als genetisches Werkzeug mit unübertroffenem Anwendungspotential in allen Domänen des Lebens kann als Meilenstein in der Geschichte der Molekularbiologie bezeichnet werden (Choi und Lee, 2016; Mojica und Montoliu, 2016). Erstmals im Jahr 1987 in *E. coli* K12 identifiziert (Ishino *et al*., 1987), wurden die in regelmäßigen Abständen unterbrochenen DR-Sequenzen des CRISPR-Arrays zunächst als genetische Marker zur

Klassifizierung bakterieller Stämme genutzt, um schließlich im Jahr 2005 mit dem Immunsystem von Bakterien und Archaeen in Zusammenhang gebracht zu werden (Mojica und Montoliu, 2016). Seitdem wurde in ~40% aller sequenzierten Prokaryoten eine Vielzahl verschiedener CRISPR-Cas Systeme identifiziert (Choi und Lee, 2016). Die Entwicklung eines programmierbaren *Genome Editing*-Werkzeuges gelang jedoch erstmals auf Basis des Typ II-A CRISPR-Cas9 Systems aus *Streptococcus pyogenes* (Jinek *et al*., 2012). Das native System weist zwei regulatorische RNAs auf, die crRNA und tracrRNA, welche die Nuklease Cas9 zu einer variablen Ziel-DNA leiten und einen Doppelstrangbruch vermitteln. Typischerweise kodiert ein CRISPR-Array mehrere Ziel-DNAs in Form von 25-40 bp großen *Protospacer*-Sequenzen. Ihr Transkript, die pre-crRNA, bedarf einer RNAse III-abhängigen Prozessierung zu ausgereiften crRNAs, die daraufhin zusammen mit der tracrRNA zu RNA-Nuklease-Komplexen assemblieren. Cas9 ist nun in der Lage, die Ziel-DNA zu binden und einen DSB zu katalysieren. CRISPR-*Editing*-Systeme basieren meist auf einer tracrRNA-crRNA-Schimäre (sgRNAs), und synthetische, ~20 bp große *Protospacer*-Sequenzen ermöglichen einen universalen Einsatz dieses Werkzeuges (Jinek *et al*., 2012) (Abb. 1A iii). Durch Mutation der Nuklease-Domänen fungiert die Cas9-Variante dCas9 zudem als DNA-Bindeprotein und ermöglicht vielfältige DNA-Modifikations- sowie Expressions-Modulationssysteme (Shalem *et al*., 2015).

In der Synthetischen Mikrobiologie basiert CRISPR-vermitteltes *Genome Editing* typischerweise auf der Einführung von DSBs. Adaptierte Protokolle wurden bereits für bakterielle Modellorganismen der Gattungen *Bacillus*, *Streptomyces*, *Escherichia* oder *Corynebacterium* entwickelt (Choi und Lee, 2016). In der Prokaryontengenetik ermöglichen CRISPR-Systeme vielfältige *Editing*-Modi, erfordern jedoch aufgrund der Diversität der Komponenten, d.h. der RNA (crRNA/tracrRNA, sgRNA), der Donor-DNA (linear/zirkulär, einzel/doppelsträngig), der Reparatur-Mechanismen (nativ/exogen, HDR/NHEJ, Größe der homologen Region) und Organismen-spezifischer Präferenzen eine aufwändige Optimierung. Die Möglichkeit zur simultanen, Selektionsmarker-freien Modifizierung mehrerer Sequenzabschnitte stellt jedoch einen essentiellen Vorteil gegenüber herkömmlichen Methoden dar. Im Falle einer Erkennung der Wirts-DNA (*self-targeting*) geht die hohe Effizienz des CRISPR-Systems außerdem mit einer hohen Letalität einher. CRISPR-basierte Selektionssysteme stellen daher eine wertvolle Ergänzung für bereits etablierte *Editing*-Ansätze dar. ssDNA *Recombineering* beispielsweise beruht auf der Elektroporation einzelsträngiger Oligonukleotide, die eine wenige bp umfassende Mutation tragen. Mit Hilfe Phagen-stämmiger DNA-Bindeproteine werden diese vermutlich während der DNA-Replikation analog zu einem Okazaki-Fragment in einem Teil der Zellen in das Genom inkorporiert (Constantino und Court, 2003). Die Identifizierung positiver Transformanten ist jedoch aufwändig. Dank CRISPR-vermitteltem *self-targeting* der nativen Zielsequenzen gelang es jedoch, den hohen *Screening*-Aufwand ssDNA

Recombineering-basierter *Engineering*-Anwendungen wie dem *multiplex automated genome engineering* (MAGE)-System erheblich zu reduzieren (Wang *et al*., 2009; Barrangou und van Pijkeren, 2016).

2.1.1.2 Rekombinasen ermöglichen umfangreiche *in-vivo* DNA-Modifikationen

HR- und NHEJ-basierte *Genome Editing*-Methoden beschränken sich meist auf lokale Mutationen von geringem Umfang. Ortsspezifische Rekombinasen ermöglichen hingegen die Deletion, Inversion und Insertion sehr großer DNA-Moleküle und haben sich daher längst in der Prokaryoten- und Eukaryotengenetik als unentbehrliche *Genome Engineering*-Werkzeuge etabliert (Wirth *et al*., 2007; Brown *et al*., 2011; Ow, 2016; Turan *et al*., 2013). Die natürlichen Funktionen ortsspezifischer Rekombinasen sind mannigfach. Sie vermitteln die Resolution der Chromosomen-Dimere nach der bakteriellen DNA-Replikation (Sherratt *et al*., 2004), katalysieren während einer Bakteriophagen-Infektion die chromosomale Integration des Lambda-Genoms in *E. coli* (Weil und Signer, 1968; Echols, 1970), oder erzeugen mittels Promotor-Inversion die flagellare Phasenvariation in *Salmonella enerica* (Haykinson *et al*., 1996). Integrations-Reaktionen erfolgen hierbei zwischen individuellen DNA-Molekülen, wohingegen Exzisionen und Inversionen auf demselben DNA-Molekül und in Abhängigkeit der relativen Orientierung Rekombinase-spezifischer Erkennungssequenzen auftreten (Abb. 2.2A).

Trotz einer Vielzahl potentieller Regulatoren und struktureller Besonderheiten der involvierten DNA-Sequenzen unterliegt der Rekombinationsmechanismus einem allgemeinen Konzept (Grindley *et al*., 2006). Demnach binden meist mehrere Rekombinase-Monomere spezifisch an zwei Erkennungssequenzen, deren *Crossover*-Regionen dabei in räumliche Nähe geraten und eine definierte Orientierung zueinander einnehmen (Abb 2.2B). Im synaptischen Komplex vermittelt die Rekombinase eine Exzision der DNA-Stränge und katalysiert zugleich deren reziproke Verknüpfung, um daraufhin die Rekombinations-Produkte wieder zu entlassen. Analog zum Typ IB Topoisomerase-Mechanismus vollzieht sich der DNA-Strangbruch im synaptischen Komplex nicht durch Hydrolyse der DNA, sondern durch einen Phosphoryl-Transfer auf eine AS-Seitenkette der Rekombinase. Bei dieser Seitenkette handelt es sich um Serin- oder Tyrosin-Reste, woraus sich ein fundamentaler Unterschied im katalytischen Mechanismus ergibt. Aus diesem Grund werden Rekombinasen entweder der Serin- oder der Tyrosin-Familie zugeordnet, und wichtige Vertreter beider Familien sind in Abbildung 2.2C aufgeführt.

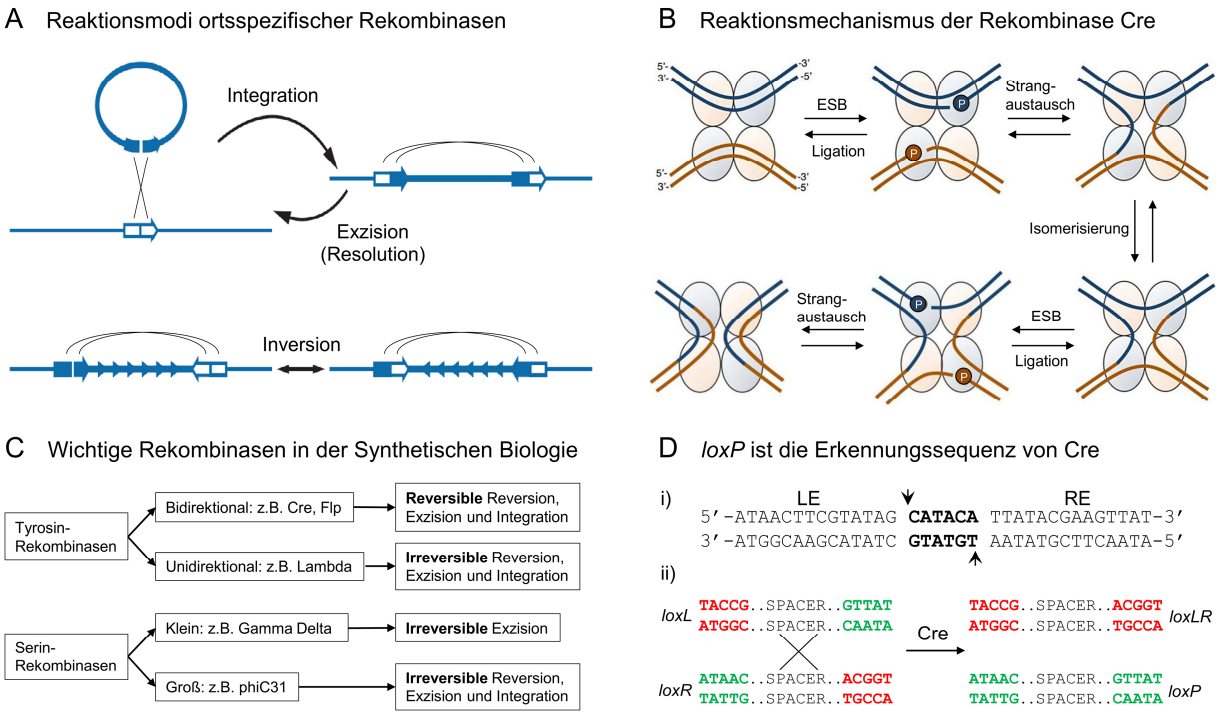


Abbildung 2.2: Funktion und Mechanismus ortsspezifischer Rekombinasen. **A:** Reaktionsmodi ortsspezifischer Rekombinasen. Abhängig von der Orientierung zweier Erkennungssequenzen (blaue Pfeile) katalysieren Rekombinasen nicht-homologe Rekombinationsereignisse, die zu einer Integration, Exzision oder Inversion der involvierten DNA-Fragmente führen können. Modifiziert nach Grindley *et al.* (2006). **B:** Darstellung des reversiblen Rekombinationsmechanismus am Beispiel der Tyrosinrekombinase Cre. Vier Einheiten des Enzyms binden zwei DNA-Moleküle und bilden eine Rekombinationssynapse (Berg *et al.*, 2013). In dieser findet eine sukzessive Einzelstrang-Spaltung der beiden DNA-Moleküle statt, wobei sich die 3'-Phosphatgruppe (P) mit einem Tyrosinrest der Rekombinase verbindet. Der Strangaustausch findet statt, indem die freie 5'-Hydroxylgruppe des reziproken Stranges die DNA-Tyrosin-Einheit angreift und eine neue Phosphodiesterbindung bildet. Dieser Prozess führt zur Bildung einer Holliday-Struktur, dem Intermediat der Rekombinationsreaktion. Nach Isomerisierung der Struktur kann sich der Austauschprozess mit den verbleibenden Strängen wiederholen, woraufhin der Komplex dissoziiert und das Rekombinationsprodukt entlassen wird (Jayaram *et al.*, 2015; Gaj *et al.*, 2014). Modifiziert nach Gaj *et al.* (2014). **C:** Beispiele verbreiteter Tyrosin- und Serin-Rekombinasesysteme. Ein grundlegendes Unterscheidungsmerkmal ist der uni- bzw. bidirektionale Rekombinationsmodus (Wang *et al.*, 2011b). **D:** Varianten der Cre-Erkennungssequenz und Rekombinationsprodukte. (i) Die native *loxP* (*locus of crossover in phage P1*)-Sequenz setzt sich aus einer zentralen, 6 bp großen *Crossover*- bzw. *Spacer*-Region und den 14 bp umfassenden, Rekombinase-bindenden Elementen LE und RE zusammen. Pfeile markieren die Cre-vermittelte Schnittstelle während einer Rekombination (Van Duyne, 2001; Turan und Bode, 2011). (ii) *lox* Sequenzen mit Mutation in der LE- (*loxL*) oder RE- (*loxR*) Region ermöglichen gerichtete Rekombinationen, da *loxLR* als Rekombinationsprodukt mit beidseitig mutierter Erkennungssequenz eine drastisch reduzierte Rekombinationsfrequenz aufweist (Albert *et al.*, 1995).

Als simultane Integrasen, Resolvasen und Invertasen stellen Cre und Flp wichtige Werkzeuge für *Genome Engineering*-Prozeduren in Alphaproteobakterien wie *S. meliloti* und *A. tumefaciens* dar (Harrison *et al.*, 2011; Milunovic *et al.*, 2014; Hu *et al.*, 2016; diCenzo *et al.*, 2014). Das Cre/*loxP*-System wurde vor über 30 Jahren im Bakteriophagen P1 identifiziert, wo es während der lysogenen Phase analog zum bakteriellen Xer/*dif*-System als Resolvase die Monomerisierung des Episoms vornimmt (Sternberg *et al.*, 1981; Abremski und Hoess, 1984). Das Flp/*FRT*-System hingegen entstammt dem 2µm-Plasmid aus *S. cerevisiae* und gewährleistet dort die Initiation der *rolling-circle*-Replikation sowie eine stabile Kopienzahl dieses extrachromosomalen Replikons (Chan *et al.*, 2013; Broach *et al.*, 1982). Beide Enzyme

teilen in Hinblick auf die Entwicklung eines simplen, flexiblen Rekombinationssystems gegenüber der Lambda-Integrase, einem dritten prominenten Vertreter aus der Familie der Tyrosin-Rekombinasen, entscheidende Vorzüge: Die Rekombination verläuft bidirektional und unabhängig von zusätzlichen Hilfsproteinen (Grindley *et al*., 2006).

Cre und Flp binden als Dimere an eine 34 bp große Erkennungssequenz, die aus einer 8 bp großen, zentralen *Crossover*- bzw. *Spacer*-Region und den 13 bp flankierenden LE- (*left element*) und RE- (*right element*) Sequenzen besteht (Turan und Bode, 2011) (Abb. 2D). Im Unterschied zu Cre verlangt Flp für eine effektive Integrationsreaktion zudem ein weiteres, 13 bp großes Segment, das ein drittes Flp Monomer bindet (Turan *et al*., 2011). Ein wichtiges Unterscheidungsmerkmal ist außerdem die Position in der *Crossover*-Region, an welcher die Rekombinase den ESB vornimmt (Abb. 2D). Während Flp den Schnitt an der -4 und +4 Position katalysiert, trennt Cre die *Spacer*-Sequenz an Position +3 und -3. Hieraus ergibt sich für die Cre/*loxP*-Reaktion eine um 2 bp reduzierte *Crossover*-Region, wodurch die Mutationskapazität des *loxP-Spacers* im Vergleich zum *FRT-Spacer* eingeschränkt wird (Turan *et al*., 2011). Die Mutation des *Crossover*-Bereichs ist jedoch eine wichtige Strategie zur Erzeugung heterospezifischer Erkennungssequenzen (*lox Spacer*-Mutanten). Diese werden von der Rekombinase als Substrat akzeptiert, rekombinieren jedoch nicht mehr mit dem nativen Derivat und im besten Fall nur mit sich selbst (Lee und Saito, 1998; Turan und Bode, 2011; Missirlis *et al*., 2006). Variable Bindungsspezifitäten der LE/RE-Elemente implizieren einen dritten, wichtigen Unterschied zwischen Cre und Flp. Während bereits geringfügige Mutationen dieser Elemente die Flp-Rekombination korrumpieren (Senecoff *et al*., 1988), akzeptiert Cre eine Mutation der äußeren vier Basenpaare der *loxP*-Sequenz. So wurden mit *lox66* (*loxR*) und *lox71* (*loxL*) LE/RE-Mutationen identifiziert (Albert *et al*., 1995), die eine effektive Cre-Rekombination erlauben, wohingegen das vollständig mutierte Rekombinationsprodukt (*loxLR*) der Rekombinase nicht mehr zur Verüfgung steht (Abb. 2D). Da der reversible Reaktionsmodus der nativen Rekombinase die Gefahr birgt, das angestrebte Rekombinationsprodukt wieder zu verlieren, können LE/RE-Mutanten insbesondere in vielschrittigen Rekombationstrategien von großem Vorteil sein.

2.1.2 *Genome Engineering* in der Synthetischen Mikrobiologie

Genome Engineering ist in dieser Arbeit als die Konzeption und Konstruktion einer umfangreichen, genomischen Modifikation unter Anwendung geeigneter *Genome Editing*-Werkzeuge definiert. Eine Facette der Synthetischen Biologie ist hierbei die Rekonstruktion bzw. *de-novo* Synthese des Erbguts (*Bottom-Up*), oder aber dessen maximale Reduktion (*Top-Down*). Das ultimative Ziel dieser komplementären Ansätze ist die Erzeugung eines Minimalsystems, das auf die Aufrechterhaltung der essentiellen Prozesse reduziert ist. Im Kontext des *Metabolic Engineering* kann dies die Identifizierung

der Kernkomponenten eines Biosynthesewegs beabsichtigen. Im *Genome Engineering* hingegen strebt der *Top-Down*-Ansatz die Erforschung der grundlegenden Prinzipien einer lebendigen Zelle an.

Die Kodierdichte bakterieller Genome wird auf 87% geschätzt (Kuo *et al*., 2009), und symbiotische Bakterien, die aufgrund ihrer Adaptation an einen spezifischen Wirt gewissermaßen zum Vorbild für extreme Genom-Reduktionen evolvierten, erreichen Genomgrößen von weniger als 150 kb (McCutcheon und Moran, 2012). Das Genom eines nativen Bakteriums bzw. potentiellen Chassis-Organismus weist folglich großes Optimierungspotential auf. Aus biotechnologischer Sicht birgt der *Top-Down*-Ansatz zudem das Potential, die Aufnahmekapazität eines Genoms zu erhöhen und die metabolische Effizienz eines Produktionsstammes bei der Synthese des gewünschten Produktes zu optimieren (Wu et al., 2016). Die fortschreitende Genom-Sequenzierung bakterieller Modellorganismen ermöglicht die systematische Deletionen zur Identifizierung nicht-essentieller Sequenzen wie Insertionselemente, endogener Rekombinasen, Phagen, oder duplizierter Gene. So konnte bereits vor über 10 Jahren das Genom von *E. coli* um 15% reduziert und gleichsam die genomische Stabilität und Elektroporationseffizienz erhöht werden (Pósfai *et al*., 2006). Heute lassen sich essentielle Gensequenzen auf Basis bioinformatischer und experimenteller Deletions-Studien vorhersagen (Gil *et al*., 2004; Baba *et al*., 2006; Forster und Curch, 2006). Der MEGA (*multiple essential genes assembling*)-Ansatz stellt eine elegante Methode zur seriellen Deletionen großer DNA-Segmente im Umfang mehrerer hundert kb dar (Xue *et al*., 2015). Hierzu werden essentielle Gene zunächst in *S. cerevisiae* mittels *transformation-associated recombination* (TAR)-Klonierung zu einem Gen-Cluster assembliert, das mit Hilfe eines *Shuttle*-Vektors in *E. coli* eingebracht wird und dort die nativen Sequenzen *in-trans* komplementiert. Dies ermöglicht die Deletion korrespondierender, chromosomaler Regionen und die anschließende Lambda-Red-vermittelte Integration des synthetischen Gen-Clusters in die Deletionsregion.
Negative Effekte auf das Zellwachstum (Kurokawa *et al*., 2016) und eine Abnahme der genomischen Plastizität solcher „*clean genome*“ Stämme (Akeno *et al*., 2014) verdeutlichen jedoch die grundsätzliche Problematik dieser Vorgehensweise. Das Design optimierter Genome erfordert nicht nur Kenntnis über die Funktion einzelner Gene, sondern auch ein umfassendes Verständnis zu übergeordneten Strukturen wie der Genom-Organisation mitsamt ihrer Implikationen auf Zellwachstum und Zellzyklusregulation (Esnault *et al*., 2007; Boccard *et al*., 2005; Xie *et al*., 2015). Die Akzeptanz gegenüber substantiellen Genom-Reduktionen scheint zudem Organismen-spezifisch zu sein. So konnte das Chromosom von *B. subtilis* um 23% reduziert werden, ohne Wachstum und Morphologie signifikant zu beeinflussen (Ara *et al*., 2007). Als robuster Chassis-Organismus erweist sich *B. sutilis* somit auch als nützliche Produktionsplattform in biotechnologischen Anwendungen (Li *et al*., 2016; Toya *et al*., 2014; Manabe *et al*., 2011).

In Anbetracht sinkender DNA-Synthesekosten und der Komplexität bakterieller Genome ist der *Bottom-Up*-Ansatz gegenüber der *Top-Down*-Strategie eine verheißungsvolle Alternative bei der Entwicklung eines Minimalgenoms. Als eines der prominentesten Beispiele dieses Ansatzes gilt zweifelsohne die Schaffung einer synthetischen Zelle auf Basis eines vollsynthetischen Chromosoms im Jahr 2010 (Gibson *et al*., 2010). Hierzu wurden über 1078 synthetische DNA-Fragmente in einer dreischrittigen Assemblierungsstrategie zu einer ~1,1 Mb großen Variante des Genoms von *Mycoplasma mycoides* fusioniert. Der finalen Transplantation des *M. mycoides* JCVI-syn1.0 Genoms in *M. capricolum* gingen unverzichtbare technische Durchbrüche voraus. Nachdem sich *E. coli* basierte Methoden zur Assemblierung subchromosomaler DNA-Fragmente als ungeeignet erwiesen und auch *in-vitro* DNA-Assemblierungstechniken längst an ihre Grenzen gestoßen waren, gelang im Jahr 2008 erstmals mit Hilfe der TAR (*transformation-associated recombination*)-Klonierungstechnik die Erzeugung eines vollsynthetischen, 583 kb großen *Mycoplasma*-Chromosoms in *S. cerevisiae* (Gibson *et al.,* 2008). Die auf der hohen Rekombinationseffizienz in *S. cerevisiae* beruhende DNA-Assemblierungsmethode, die ursprünglich zur Klonierung humaner DNA entwickelt wurde (Larionov *et al*., 1996), ermöglicht die gerichtete Fusion mehrerer DNA-Fragmente mittels *End Joining* in einen geeigneten Ziel-Vektor (Kouprina und Larionov, 2016). Die TAR-Klonierung stellt bis heute die effizienteste Technik dar, um artifizielle Konstrukte aus bis zu 300 kb großen DNA-Fragmenten eukaryotischen und prokaryotischen Ursprungs zu klonieren. Als Alternative zur Genom-Synthese via DNA-Assemblierung wurde zudem der Transfer bakterieller Genome in *S. cerevisiae*-stämmige Spheroplasten etabliert (Karas *et al*., 2013), womit die *Genome Editing*-Werkzeuge der Hefe-Plattform grundsätzlich auch zur genomischen Modifizierung genetisch unzugänglicher Organismen zur Verfügung stehen. Um den *Engineering*-Kreislauf zwischen *S. cerevisiae* und *Mycoplasma* zu schließen, wurden zudem Transplantations-Methoden zum Transfer isolierter Chromosome nach *Mycoplasma* ausgearbeitet (Lartigue *et al*., 2007; Lartigue *et al*., 2009). Dies alles war wegbereitend für die Entwicklung eines auf JCVI-syn1.0 basierenden Minimalgenoms, das heute mit 473 Genen das kleinste bekannte Chromosom einer autonom replizierenden Zelle darstellt (Hutchison *et al*., 2016).

Insbesondere *in-vitro*-Methoden wie die Isolierungen subchromosomaler DNA-Intermediate und deren Transformation minimieren jedoch die Effizienz ambitionierter *Engineering*-Strategien. Um die mit der TAR-basierten Genom-Assemblierung assoziierten Probleme zu lösen, wurde zuletzt die *Cas9-facilitated homologous recombination assembly* (CasHRA)-Methode entwickelt (Zhou *et al*., 2016). Hierbei werden zirkuläre Donor-Vektoren, die die zu assemblierenden Fragmente tragen, zunächst mittels Sphäroplasten-Fusion in *S. cerevisiae* vereint. Erst dann generiert Cas9 *in-vivo* die lineare DNA, die nach dem Prinzip der TAR-Klonierung über homologe Flanken fusioniert. Da die Aufreinigung und Co-Transformation

großer DNA-Intermediate entfällt, gelang die Assemblierung eines 1,03 Mb großen *E. coli* Minimalgenoms mit hoher Effizienz (60-80%). Tatsächlich stellt der *in-vitro* Transfer des assemblierten Chromosoms von *S. cerevisiae* nach *E. coli* eine technische Hürde dar, die es noch zu überwinden gilt.

Bei *B. subtilis* handelt es sich um die einzig verfügbare *Engineering*-Plattform auf Basis eines Prokaryoten, die eine mit *S. cerevisiae* vergleichbare Klonierungskapazität aufweist. Entscheidend ist hierbei die Fähigkeit des Chromosoms, bis zu 3,5 Mb große DNA-Fragmente heterologen Ursprungs stabil zu propagieren (Itaya *et al*., 2005). Die natürliche Kompetenz dieses Organismus geht zudem mit der Eigenschaft einher, lineare DNA in hoher Effizienz in das Chromosom zu integrieren (Dubnau, 1991). Das *Bacillus subtilis genome* (BGM)-Vektorsystem macht sich diese Eigenschaften zunutze (Itaya, 1999) und ist die Grundlage der „Domino-Klonierung", die zur Assemblierung rekombinanter Genome konzipiert wurde (Itaya *et al*., 2008). Analog zur TAR-basierten Assemblierung beruht das System auf einer sukzessiven Fusion linearisierter DNA-Fragmente, die nach Transformation in *B. subtilis* via überlappender, homologer Regionen in den BGM-Vektor integrieren. Der einschrittige Assemblierungsmodus impliziert jedoch eine vergleichsweise aufwändige Prozedur. Die als *Bacillus recombinatorial transfer* (BReT) bezeichnete *in-vivo* Klonierungsmethode ermöglicht schließlich die Isolierung des Assemblierungsprodukts (Kaneko *et al*., 2005). Aufgrund der *Bacillus*-spezifischen Eigenheit, chromosomale DNA via homologer Flanken auf ein linearisiertes Vektor-*Backbone* zu überführen, ist es möglich, assemblierte Konstrukte auf ein ektopisches Replikon zu duplizieren (Itaya und Tanaka, 1997a). Die Auslagerung auf einen konventionellen BAC-Vektor ermöglichte so den Transfer bis zu 350 kb großer DNA-Fragmente nach *E. coli*.

Genetic Engineering setzt einige biologische und technische Voraussetzungen voraus, die sich von etablierten Chassis-Organismen ableiten lassen. Klonierungsstrategien werden oft durch eine mangelhafte Transformationskapazität des Rezipienten limitiert. Eine hohe Transformationseffizienz ist jedoch unabdingbar, um den Transfer sehr großer DNA-Moleküle bzw. aufwändige Klonierungsprozesse in angemessener Zeit zu bewerkstelligen. Dabei gilt es mehrere Hindernisse zu überwinden. Durch Manipulation der Beschaffenheit von Zellwand und –membran konnte die Transformationsbarriere beispielsweise in *B. subtilis* mit Hilfe von Sorbitol und Mannitol (Xue *et al*., 1999), in *S. cerevisiae* mit PEG (Karas *et al*., 2013), und in *Corynebacterium glutamicum* mit Glycin (Ruan *et al*., 2015) reduziert werden. In der Natur haben sich zudem effektive Abwehr-Mechanismen entwickelt, die den Prozess der DNA-Aufnahme behindern. Neben dem CRISPR-System stellen Restriktions-Modifikations (RM)-Systeme einen effektiven Mechanismus dar, artfremde DNA zu degradieren (Vasu *et al*., 2012). RM-Systeme des Typs I beispielsweise sind multifunktionale Enzyme, deren Untereinheiten HsdMS DNA-

Methylaseaktivität aufweisen, während die Restriktionsendonuklease HsdR die unmodifizierten Erkennungssequenzen der Fremd-DNA attackiert (Murray *et al*. 2000). Tatsächlich war in wichtigen bakteriellen Chassis-Organismen wie *M. mycoides* (Lartigue *et al*., 2009; Karas *et al*., 2013), *B. subtilis* (Itaya und Tanaka, 1997b), *Staphylococcus aureus* (Sadykov 2016), oder *Streptomyces griseus* (Suzuki *et al*., 2011) eine Inaktivierung des RM-Systems vonnöten, um umfangreiche *Engineering*-Prozeduren zu realisieren.

Mit zunehmender Molekülgröße gewinnen zudem DNA-Vehikel an Bedeutung, die einen flexiblen DNA-Austausch zwischen verschiedenen Spezies ermöglichen. Insbesondere in ambitionierten Assemblierungs-Projekten ist der DNA-Transfer zwischen *Engineering*- und Ziel-Organismus ein limitierender Faktor, was letztlich auf den hohen technischen Anspruch zugrundeliegender *in-vitro*-Techniken und Limitierungen klassischer Transformationsmethoden zurückgeführt werden kann. Typ IV Sekretionssysteme etablieren hingegen natürliche DNA-Transfersysteme, die einen effizienten Gentransfer zwischen einer Vielzahl Gram-negativer und auch Gram-positiver Bakterien ermöglichen (Chandran 2013). Der RK2/RP4-basierte Konjugationsapparat stellt daher eine zweckdienliche Alternative dar, um mit Hilfe mobilisierbarer *Shuttle*-Vektoren den DNA-Transfer zwischen *E. coli* und kompatiblen Spezies zu bewerkstelligen (Kvitko *et al*., 2013). *Genetic Engineering*-Anwendungen profitieren zudem von DNA-Vehikeln, die entsprechend *S. cerevisiae*-kompatiblen oder BGM-Vektoren eine stabile Vererbung selbst mehrerer Mb großer DNA-Moleküle gewährleisten (Gibson *et al*., 2010; Kaneko *et al*., 2005) und zugleich deren Modifizierung *in-vivo* erlauben.

2.2 *Sinorhizobium meliloti* ist ein Stickstoff-Fixierer mit landwirtschaftlicher Bedeutung

Rhizobien sind bodenbewohnende Mikroorganismen aus der Klasse der Alphaproteobakterien, die für ihre endosymbiotische Stickstofffixierung in Wurzelknöllchen von Leguminosen wie der Sojabohne (*Glycine sp.*), Erbse (*Pisum sp.*), oder Luzerne (*Medicago sp.*) bekannt sind (Gibson 2008; Jones *et al*., 2007). *S. meliloti* ist ein prominenter Vertreter und Modellorganismus für diese mutualistische Bakterien-Pflanzen-Interaktionen. Zur Einrichtung der Lebensgemeinschaft sondern potentielle Wirtspflanzen (u.a. *Medicago truncatula* und *Medicago sativa*) spezifische Signalstoffe über ihre Wurzelexudate in den Wurzelbereich (Rhizosphäre) aus. Flavonoide wie Luteolin spielen in dieser frühen Etablierungsphase eine zentrale Rolle und induzieren die Synthese und den Export so genannter Nodulations-Faktoren (Geurts *et al*., 2005). Ein wechselseitiger Signalaustausch induziert daraufhin die Bildung des Knöllchenmeristems an der Leguminosenwurzel (Gibson *et al*., 2008). Außerdem führt er zu einer chemotaktischen Anlagerung der freilebenden Bakterien an die Wurzelhaare der Wirtspflanze, welche die Rhizobien durch Krümmung einschließen und diesen via Infektionsschlauch Zugang zu dem neugebildeten, kortikalen Meristem gewähren (Esseling *et al*., 2003). In diesen als Wurzelknöllchen

bezeichneten Strukturen findet eine Einkapselung von *S. meliloti* sowie dessen Differenzierung zu vergrößerten und verzweigten Zellen statt. In der Stickstofffixierungszone eines Wurzelknöllchens reduziert schließlich der sauerstoffempfindliche Nitrogenasekomplex atmosphärischen Stickstoff (N_2) zu Ammoniak, der unmittelbar zu Glutamin und Asparagin assimiliert und der Pflanze im Austausch gegen C4-Metabolite zur Verfügung gestellt wird (Mergaert *et al*., 2006).

Rhizobien spielen in der modernen Landwirtschaft eine fundamentale Rolle. Es wird geschätzt, dass durch menschliche Prozesse pro Jahr rund $120*10^9$ kg N_2 in biogene Stickstoffverbindungen konvertiert werden, während die rhizobielle Stickstofffixierung mit landwirtschaftlich relevanten Nutz- und Futterpflanzen auf $33\text{-}46*10^9$ kg N_2 taxiert wird (Rockström *et al*., 2009; Herridge *et al*., 2008). Aufgrund des hohen Energieaufwandes der Synthese von molekularem Wasserstoff, einem Substrat des Haber-Bosch-Verfahrens zur chemischen Erzeugung von Ammoniak, partizipiert die industrielle Herstellung künstlicher Stickstoffdünger mit ~1,2% in nennenswertem Ausmaß am Weltenergiebedarf (Tallaksen *et al*., 2015). Die genetische Rekonstruktion der biologischen Stickstofffixierung und ihre Übertragung auf weitere Kulturpflanzen sind daher zu einem zentralen Thema der Synthetischen Biologie avanciert (Mus *et al*., 2016).

2.2.1 Alphaproteobakterien besitzen mehrteilige Genome mit *repABC*-basierten Replikationsursprüngen

Die organismische und genetische Diversität innerhalb der Alphaproteobakterien ist enorm (Ettema und Andersson, 2009). Der Klasse gehören phototrophe (*Rhodobacter* sp.) und aquatisch lebende (*Caulobacter* sp.) Spezies an, aber auch Tier- und Humanpathogene (*Brucella* sp., *Bartonella* sp., *Rickettsia* sp.), Pflanzenpathogene (*Agrobacterium sp*.) und Pflanzensymbionten (*Sinorhizobium* sp., *Rhizobium* sp., *Bradyrhizobium* sp.), deren Spezialisierung an ihren jeweiligen Wirtsorganismus nur das Extrem einer bemerkenswerten Adaptationsfähigkeit an eine Vielzahl verschiedener Lebensräume darstellt.

Das hohe Maß an Diversität geht mit einer beträchtlichen Heterogenität bezüglich der Genomgröße und -architektur einher (Sallström und Andersson, 2005; Harrison *et al*., 2010; Moreno, 1998), die den horizontalen Gentransfer und damit verbundenen hohen Differenzierungsgrad dieser Organismen begünstigen (Pini *et al*., 2011; Galardini *et al*., 2013). Insbesonders schnell wachsende Vertreter der Rhizobien zeichnen sich durch ein mehrteiliges Genom aus, das sich aus einem primären Chromosom und bis zu sechs Megaplasmiden zusammensetzen kann (Mazur und Koper, 2012; González *et al*., 2006). Letztere enthalten vornehmlich adaptive Gene und umfassen den oft umfangreichen, akzessorischen Teil des Genoms (Galardini *et al*., 2013). Im Laufe der evolutionären Anpassung gleicht sich der GC-Gehalt von Chromosom und Megaplasmid an, und auch essentielle Gene werden auf das extrachromosomale

Replikon verlagert (Harrison *et al*., 2010). Megaplasmide entwickeln sich dabei zu Chromiden, die als sekundäre Chromosome verstanden werden können und die Gültigkeit der scharfen Abgrenzung (del Solar *et al*., 1998) zwischen Chromosom und Plasmid infrage stellen.

S. meliloti besitzt neben einem Chromosom (3,65 Mb) ein 1,35 Mb großes Megaplasmid (pSymA) und ein 1,68 Mb großes Chromid (pSymB) (Galibert *et al*., 2001) (Abb. 2.3A). pSymA und pSymB gehören der *repABC*-Plasmidfamilie an und sind damit Teil der *repC*-Superfamilie, die nahezu alle extrachromosomalen Replikons rhizobieller Spezies umfasst (Cevallos *et al*., 2008). Der DNA-Replikationsinitiator RepC ist hierbei namensgebend und exklusiv in Alphaproteobakterien bekannt. *repABC*-basierte Replikons weisen eine erweiterte Transkriptionseinheit auf, die neben RepC einen zum ParAB-System homologen Segregationsapparat kodiert. Abb. 2.3B stellt den konservierten Aufbau eines *repABC*-basierten Replikationsursprungs dar, der im Folgenden als Einheit aus Operon und assoziierter Elemente definiert sei.

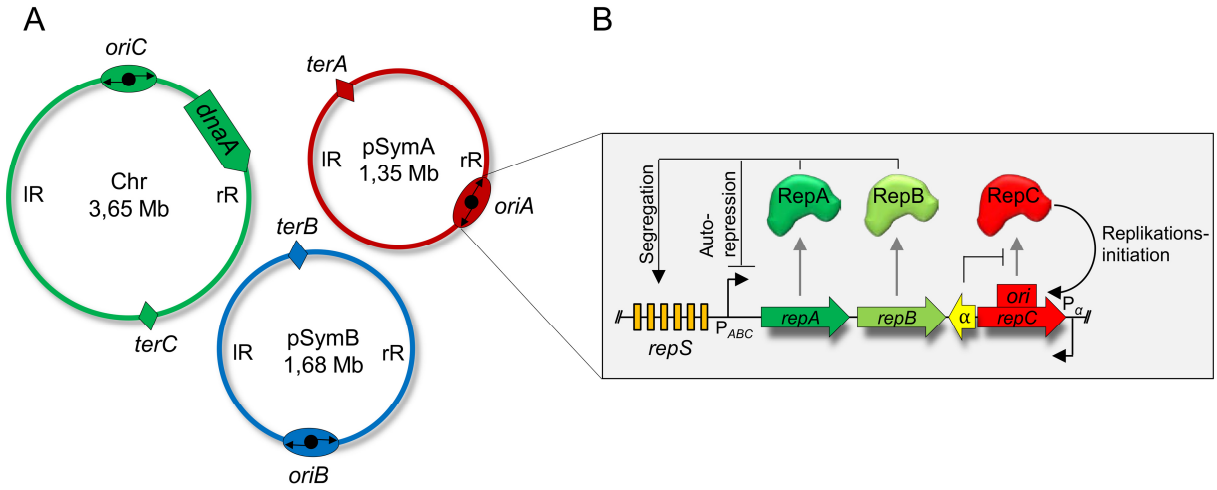


Abbildung 2.3: Organisation des mehrteiligen Genoms von *S. meliloti*. **A:** Architektur der Replikons. Struktur-verleihende Elemente geben Chromosom (Chr), Megaplasmid pSymA und Chromid pSymB eine symmetrische Struktur, die die Definition eines linken und rechten Replichors (lR, rR) mit *ori-ter*-Orientierung ermöglicht (Hendrickson und Lawrence, 2006). Das Gen des chromosomalen Replikationsinitiators DnaA liegt mit ~400 kb Abstand ungewöhnlich weit vom chromosomalen Replikationsursprung (*oriC*) entfernt (Galibert *et al*., 2001). *oriA/B*: Replikationsursprung pSymA/B. *terA/B/C*: Terminus-Region von pSymA/B und Chromosom. **B:** Schematische Darstellung des konservierten Aufbaus eines *repABC*-basierten Replikationsursprungs. Das *repS*-Motiv ist eine 16 bp große, palindromische RepB-Bindesequenz, die in ein- bis sechsfacher Ausführung über die *repABC*-Region verteilt sein kann (Soberón *et al*., 2004; MacLellan *et al*., 2006) und die RepAB-basierte Segregation vermittelt (Pinto *et al*., 2012). RepB stimuliert zudem die Bindung von RepA an einen Operator innerhalb des Promotors (P_{ABC}), wodurch eine Reprimierung der Transkription des Operons erzielt wird. Eine Regulation der *repC* Expression erfolgt außerdem auf post-transkriptioneller Ebene mittels der regulatorischen ctRNA (α), die grundsätzlich zwischen *repB* und *repC* in reverser Orientierung kodiert ist. Die ctRNA bindet spezifisch die SD-Seuqenz der RepC mRNA, wodurch die Translationsrate des Replikationsinitiators reguliert wird (Cervantes-Rivera *et al*., 2010; Rivera-Urbalejo *et al*., 2015). P_{α}: Promotor der ctRNA.

Wie bereits im Jahr 1989 anhand des Tumor-induzierenden Plasmids pTi aus *A. tumefaciens* nachgewiesen wurde (Tabata *et al*., 1989), enthält die *repABC*-Kassette alle zur Vererbung des Replikons

notwendigen Elemente (Ramírez-Romero *et al*., 2000; Pappas und Winans, 2003b). *repC* stellt hierbei die Minimalregion des Replikationsursprungs dar (Cervantes-Rivera *et al*., 2011; Cevallos *et al*., 2008). Die Gensequenz weist nahe des Zentrums eine ~150 bp große, AT-reiche Region auf, die RepC bindet und vermutlich im Zusammenspiel mit noch unbekannten Faktoren die Assemblierung des Replikationskomplexes vermittelt (Pinto *et al*., 2011). AT-reiche Sequenzen sind typische Bindestellen für chromosomale Replikationsinitiatoren wie DnaA (Rajewska *et al*., 2012), das auch in *S. meliloti* für die Replikationsinitiation des Chromosoms verantwortlich ist (Sibley *et al*., 2006). DnaA konnte jedoch keine regulatorische Funktion zur Megaplasmid-Replikation nachgewiesen werden (Frage *et al*., 2016), und Faktoren, die diesen Prozess mit dem Zellzyklus synchronisieren, sind gänzlich unbekannt.

Zentromer-ähnliche *repS*-Sequenzen vermitteln die Bildung eines Nukleoproteinkomplexes mit RepAB, der die Polymerisierung der ATPase RepA induziert und die treibende Kraft zur Segregation der duplizierten Replikationsursprünge bereitstellt (Gerdes *et al*., 2010; Schumacher, 2012). pSymA und pSymB weisen dabei in *S. meliloti* eine auf das Chromosom abgestimmte Koordination auf (Frage *et al*., 2016). Während die Replikationursprünge aller drei Replikons zu Beginn eines Zellzyklus nahe des alten Zellpols lokalisieren, wandern diese Regionen zunächst in Richtung Zellmitte, woraufhin ihre replizierten Kopien in definierter Reihenfolge zu beiden Zellpolen segregieren. Auch die molekularen Mechanismen, die dieses räumlich und zeitlich regulierte Segregationsverhalten koordinieren, sind noch unverstanden.

Eine stabile Vererbung und sehr niedrige Kopienzahl sind Chromosomen-ähnliche Eigenschaften *repABC*-basierter Plasmide und auf die feinjustierte Expressionsregulation des zugrundeliegenden Operons zurückzuführen (Pinto *et al*., 2012; Mazur und Koper, 2012; Cevallos *et al*., 2008). Die Merkmale sind außerdem auf artfremde Plasmide und über die Artgrenze hinweg übertragbar. So weist in *S. meliloti* ein mit $repAB_{pSymA}$-Kassette versehener Vektor in einem pSymA-freien Hintergrund eine stabilisierte Vererbung auf (MacLellan *et al*., 2006). Ähnliches wurde auf Basis der RtTA1-stämmigen *repAB*- bzw. *repABC*-Regionen aus *Rhizobium leguminosarum* bv. *trifolii* dokumentiert, die die stabile Segregation rekombinanter Plasmide in *A. tumefaciens* vermitteln (Koper *et al*., 2016). Mit *repS*, *repAB*, der *ctRNA* und *repC* wurden zudem alle bekannten Komponenten des Replikationsursprungs als *in-trans* wirkende Inkompatibilitätsfaktoren identifiziert (Pérez-Oseguera und Cevallos, 2013; Cervantes-Rivera *et al*., 2011; Ramírez-Romero *et al*., 2000; Yip *et al*., 2015; Rivera-Urbalejo *et al*., 2015; Koper *et al*., 2016; MacLellan *et al*., 2006; Mazur *et al*., 2011), die eine Co-Existenz mehrerer Kopien desselben Derivates unterbinden und gleichsam die Existenz einer Vielzahl verschiedener Inkompatibilitätsgruppen implizieren. Ein ausdifferenziertes Kompatibilitätssystem sowie die autonome Funktionalität sind bemerkenswerte Eigenschaften bislang charakterisierter *repABC*-Kassetten und könnten zur Entwicklung neuer Vektorsysteme beitragen, die sich die Vorzüge dieser Replikationsursprünge zunutze machen und

den Bedarf neuer DNA-Vehikel für synthetisch-biologische Anwendungen in *S. meliloti* und verwandten Spezies bedienen (Yip *et al*., 2015).

2.2.2 *Sinorhizobium meliloti* als Modellorganismus der Synthetischen Mikrobiologie

Genetische Modifikationen in *S. meliloti* beruhen üblicherweise auf der Konjugation von DNA-Konstrukten, die mit Hilfe replizierender, pBBR1- und RK2-basierter *broad-host-range* Vektoren, bzw. integrativer, pMB1-basierter Plasmide in die Zelle eingebracht werden (Khan *et al*., 2008; Hübner *et al*., 1991, Schäfer *et al*., 1994). Nicht-replikative Konstrukte integrieren in *S. meliloti* via HR spezifisch in einen genomischen Lokus, der durch eine homologe DNA-Sequenz definiert wird. Zur anschließenden Deletion des Plasmid-*Backbones*, die mittels eines zweiten HR-Ereignisses erzielt werden kann, wurde ein Gegenselektions-Verfahren zur Identifizierung positiver Transkonjuganten entwickelt (Hynes *et al*., 1989; Schäfer *et al*., 1994). Bei diesem Marker handelt es sich um *sacB*, dessen Produkt, die Levansucrase, das für die Zellen toxische Polysaccharid Levan synthetisiert. Die „Saccharose-Selektion" erlaubt eine positive Selektion auf die Deletion eines definierten Sequenzabschnittes (oder Plasmids), da der hiermit einhergehende Verlust des *sacB*-Gens der Zelle ermöglicht, auf Saccharose-haltigem Medium zu wachsen.

Umfangreiches *Genetic Engineering* setzt nicht nur die Verfügbarkeit adäquater DNA-Vehikel und genetischer Werkzeuge voraus, es profitiert auch von einer hohen genomischen Plastizität des Chassis-Organismus (Kap. 2.1.2). Tatsächlich weist die mehrteilige Genomarchitektur rhizobieller Arten diesbezüglich eine erstaunliche Flexibilität und Robustheit selbst gegenüber Co-Integrationen ganzer Replikons auf (Mavingui *et al*., 2002; Guo *et al*., 2003). *S. meliloti* toleriert außerdem substantielle Genom-Reduktionen. Das Tn5-Transposon ist das erste Werkzeug, das umfangreiche *Genome Engineering*-Ansätze dieser Art ermöglichte (Hynes *et al*., 1989). In seriellen Transposon-Mutagenese-Runden und unter Zuhilfenahme des Gegenselektionsmarkers *sacB* gelangen so unspezifische Deletionen viele hundert kb umfassender Regionen und sogar die vollständige Deletion von pSymA (Hynes *et al*., 1989; Oresnik *et al*., 2000). Dank der nachfolgenden Identifizierung und partiellen Analyse von 211 Toxin-Antitoxinsystemen, der Identifizierung der pSymB-kodierten, essentiellen Gene *engA* und $tRNA^{arg}$, sowie der Etablierung Flp/*FRT*- und phiC31-basierter Deletions- bzw. Integrationssysteme (Kap. 2.1.1.2) wurde der *Top-Down*-Ansatz in *S. meliloti* jedoch drastisch optimiert (Makarova *et al*., 2009; Milunovic *et al*., 2014; diCenzo *et al*., 2013). So konnte mit der Deletion beider Megaplasmide die bislang größte genomische Reduktion eines prokaryotischen Genoms erzielt und auf Basis dieses Minimalgenoms eine *Engineering*-Plattform zur genetischen Rekonstruktion der symbiotischen Stickstoffixierung generiert werden (diCenzo *et al*., 2014; diCenzo *et al*., 2016). *S. meliloti* könnte sich darüber hinaus als Produktionsstamm für Exopolysaccharide und Vitamine eignen (Becker, 2015; Dong *et al*., 2016).

In Anbetracht des breiten Methodenspektrums, welches den eindrücklichen *Genetic Engineering*-Anwendungen in bekannten Chassis-Organismen zugrunde liegt (Kap. 2.1.2), wird jedoch deutlich, dass vergleichbare Verfahren in *S. meliloti* nicht zuletzt aufgrund unzulänglicher genetischer Techniken und einer geringen Auswahl adäquater DNA-Vehikel limitiert sind. Insbesondere die HR-basierte DNA-Modifikation mittels klassischer Klonierungsvektoren zieht im Vergleich zu modernen, Vektor-unabhängigen *Editing*-Methoden (Kap. 2.1.2) zeitaufwändige Klonierungsprozesse nach sich. Bislang verfügbare Rekombinasesysteme erlauben zudem nur schrittweise Reaktionen, da im Genom verbleibende, funktionelle Elemente wie Rekombinase-Erkennungssequenzen Folgereaktionen korrumpieren. In Chassis-Organismen wie *B. subtilis* und *E. coli* kann zudem auf DNA-Vehikel zurückgegriffen werden, die die Assemblierung, Aufrechterhaltung und den Transfer sehr großer DNA-Moleküle ermöglichen (Kap. 2.1.2). Dementsprechend sind nicht nur leistungsfähigere Klonierungstechniken, sondern auch DNA-Vehikel mit hoher Klonierungskapazität notwendig, um das Potential von *S. meliloti* in biotechnologisch-synthetischen Anwendungsgebieten auszuschöpfen.

2.3 Ziele der Arbeit

S. meliloti ist ein biotechnologisch relevanter Modellorganismus mit mehrteiliger Genomarchitektur, die großes Potential hinsichtlich neuer *Genetic Engineering*-Anwendungen birgt. Das Ziel dieser Arbeit ist es daher, *S. meliloti* als Chassis-Organismus in der Synthetischen Mikrobiologie zu etablieren. Eine geringe Elektroporationskapazität erschwert jedoch die genetische Modifikationen des Wildtyp-Stammes. Aus diesem Grund soll eine *hsdR*-Deletionsmutante hergestellt und die Transformationskapazität quantifiziert werden. Unzulängliche *Genome Editing*-Werkzeuge erfordern außerdem die Entwicklung neuer genetischer Methoden. Daher soll eine effiziente *Genome Editing*-Methode zur genomischen Integration heterologer DNA gefunden, und leistungsfähige, Cre/*lox*-basierte Rekombinasesysteme zur Deletionen, Inversionen und Fusionen großer DNA-Moleküle etabliert werden. Bei der Entwicklung neuer Rekombinationssysteme stellt zudem die Optimierung der zugrundeliegenden Klonierungsstrategien sowie die Konstruktion funktioneller Elemente genetischer Schaltkreise einen zentralen Aspekt dar.

Die funktionelle Charakterisierung *repABC*-basierter Replikationsursprünge verschiedener α-Rhizobien bildet weiterhin die Grundlage einer neuen *Shuttle*-Vektor-Familie. Die Vektoren sollen in *E. coli* und in *S. meliloti* replikativ sein und klassische Klonierungen, aber auch eine Adaptation in moderne Assemblierungssysteme ermöglichen. Ein modulares Assemblierungssystem soll daher die effiziente *in-vitro* Konstruktion neuer Vektor-Derivate gewährleisten und die Ansprüche an moderne Plasmid-Systeme erfüllen. Funktionelle Einheiten eines Vektors wie Replikationsursprung, Resistenzkassette oder *multiple cloning site* (MCS) werden daher spezifischen Modulen zugeordnet und die jeweiligen „Bausteine“ in Form einer DNA-Bibliothek bereitgestellt.

In dieser Arbeit entwickelte Cre/*lox*-Rekombinationssysteme sollen auch experimentelle Ansätze zur mechanistischen Untersuchung des mehrteiligen Genoms von *S. meliloti* ermöglichen und so zu einem vertieften Verständnis der Chromosomenbiologie dieses Organismus beitragen. *repABC*-basierte Replikationsursprünge dienen hingegen der Konstruktion von DNA-Vehikeln, die in einem neuen Klonierungsverfahren die biotechnologischen Ressourcen von *S. meliloti* und verwandten Spezies genetisch zugänglich machen sollen.

3. Ergebnisse und Diskussion

3.1 Optimierung und Entwicklung neuer *Genome Editing*- und *Engineering*-Verfahren für *Sinorhizobium meliloti*

3.1.1 Etablierung von *Sinorhizobium meliloti* Rm1021 als *Engineering*-Plattform

Eine grundlegende Voraussetzung eines Chassis-Organismus in der Synthetischen Mikrobiologie ist seine genetische Zugänglichkeit, die oftmals durch eine geringe Aufnahmekapazität von Fremd-DNA limitiert wird (Kap. 2.1.2). RM-Systeme degradieren heterologe DNA mit ungeeignetem Methylierungsmuster (Murray, 2000) und reduzieren die Transformationseffizienz des *S. meliloti* Wildtyp-Stammes Rm1021 drastisch (Ferri *et al*., 2010). Um diese Barriere zu beheben, wurde eine stabile *hsdR*-Deletionsmutante hergestellt und charakterisiert (Kap. 6.3). *S. meliloti* SmΔhsdR weist bei Transformation mit einem replizierenden Vektor im Vergleich zum Wildtyp-Stamm eine ~2,6-fach erhöhte Konjugationseffizienz, und eine >1000-fach erhöhte Elektroporationseffizienz auf. Das durch *hsdMSR* kodierte RM-System stellt in *S. meliloti* folglich eine starke Transformations-Barriere dar und limitiert Anwendungen, die einer hohen Transformationseffizienz oder dem Transfer besonders großer DNA-Moleküle bedürfen. Auf den Transformationserfolg integrativer Konstrukte hat das native Restriktionsenzym hingegen keine Auswirkungen.

Ein alternativer Ansatz zur Erhöhung der Transformationseffizienz von Wildtyp *S. meliloti* ist die Modifizierung des Methylierungsmusters der Plasmid-DNA im Donor-Organismus. Zu diesem Zweck wurde im Rahmen dieser Arbeit der *E. coli*-Stamm DH5α109 hergestellt, der die *S. meliloti*-stämmigen DNA-Methyltransferasegene *ccrM* und *hsdMS* auf einem Helferplasmid unter Kontrolle des Arabinose-induzierbaren Promotors PBAD (Guzman *et al*., 1995) trägt und als Plasmid-Donor eine zehnfach erhöhte Elektroporationseffizienz in *S. meliloti* Rm1021 erzielt. Tatsächlich ist eine zusätzliche Induzierung des Systems nicht notwendig und die Hintergrundexpression des eingebrachten Operons genügt, um die maximale Effizienzsteigerung zu erzielen. Aufgrund der Größe des Helferplasmids von ~25 kb wird dessen Co-Transformation zudem effektiv unterbunden, und seine Anwendung stellt eine ähnlich effiziente, aber weniger aufwendige Alternative zu *in-vitro* Methylierungssystemen dar (Chen *et al*., 2008; Donahue *et al*., 2000; Kwak *et al*., 2002). Das Konzept der Manipulation der DNA-Methylierung im DNA-Donor birgt jedoch auch potentielle Nachteile. Wie im Falle von HsdMS sind die Ziel-Sequenzen der RM-Systeme oft unbekannt, und auch die Spezifität des heterolog exprimierten Methylasekomplexes ist fraglich. Da eine Störung des *E. coli* Methylierungssystems DNA-Reparaturmechanismen beeinträchtigen kann (Marinus, 2010), sollte insbesondere eine erhöhte Mutationsrate der Plasmid-DNA ausgeschlossen werden.

Genome Engineering setzt in der Regel eine aufwändige Präparation des Zielgenoms mit verschiedenen funktionellen Elementen voraus und profitiert daher in besonderem Maße von verbesserten *Genome Editing*-Methoden (Kap. 2.1.2). Bei seriellen genomischen Integrationen provozieren jedoch homologe Bereiche der integrativen Vektoren unerwünschte Co-Integrationen und erfordern aufwändige Klonierungsprozesse, die dies verhindern. Die Elektroporation linearer Donor-DNA mit homologen Enden, die via doppelter HR ortsspezifisch integriert, stellt daher in Modellorganismen wie *B. subtilis* oder *Pseudomonas aeruginosa* eine elegante Methode dar, um unerwünschten Co-Integrationen über Vektor-stämmige Sequenzen vorzubeugen (Itaya und Tanaka, 1990; Chen und Leung, 2016). Wie viele Prokaryoten, die keine natürliche Kompetenz aufweisen, lässt sich *S. meliloti* jedoch nicht mit linearer DNA transformieren. Um ein Plasmid-*Backbone* nach genomischer Integration im Zuge einer weiteren HR zu deletieren, wurde die s*acB*-basierte Gegenselektion auf Saccharose-Sensitivität etabliert (Kap. 2.2.2). Die Herstellung entsprechender Konstrukte sowie die zugrundeliegende Selektionsprozedur sind jedoch sehr aufwändig und limitieren die Anzahl genomischer Manipulationen in einem zeitlich annehmbaren Rahmen.

Tatsächlich führt in *S. meliloti* Rm1021 auch eine Elektroporation zirkularisierter DNA-Fragmente, bestehend aus einer homologen Sequenz und einer Resistenzkassette, zu einer genomischen Integration (Kap. 6.3). In Folge dieser Erkenntnis wurde in dieser Arbeit eine HR-basierte *Genome Editing*-Methode entwickelt (*cloning-free genome editing*, CFGE), die auf der Elektroporation zirkularisierter PCR-Produkte beruht. Hierzu werden mindestens ein zur Zielsequenz homologes DNA-Fragment und eine Antibiotikaresistenzkassette zu einem zirkulären DNA-Molekül fusioniert. PCR-Amplifikate können sowohl *blunt-end* (ungerichtet, auf zwei DNA-Fragmente beschränkt), als auch in vorgegebener Orientierung und größerer Zahl via *ligase cycling reaction* (LCR) miteinander fusioniert und anschließend in elektrokompetente Zellen eingebracht werden. Die Notwendigkeit einer homologen DNA-Sequenz sowie die mit integrativen Konstrukten vergleichbar niedrige Integrationsfrequenz legen dabei nahe, dass sich die genomische Integration mittels eines HR-basierten Mechanismus vollzieht.

CFGE wurde in dieser Arbeit erfolgreich in den Alphaproteobakterien *S. meliloti* und *A. tumefaciens*, sowie dem biotechnologisch relevanten Gammaproteobakterium *Xanthomonas campestris* zur Gen-Inaktivierung angewandt. Dies deutet auf ein potentiell breites Anwendungsgebiet der CFGE-Methode innerhalb der Proteobakterien hin. Da dank der effizienten *in-vitro* Assemblierungsmethode die aufwändige Klonierung integrativer Konstrukte entfällt, stellt CFGE die schnellste beschriebene *Genome Editing*-Technik in hiesigen Organismen dar.

Die geringe Transformationseffizienz integrativer Konstrukte schränkt CFGE-basierte Anwendungen in *S. meliloti* jedoch ein und limitiert langfristig die Attraktivität dieses Organismus in der Synthetischen

Mikrobiologie. Ein Lösungsansatz zur Steigerung der Integrationsfrequenz heterologer DNA während des Transformationsprozesses könnte in der Modulation des HR-Apparates liegen. Die chromosomale Integration heterologer DNA ist in der Regel RecA-abhängig (Kap. 2.1.1.1). Eine Inaktivierung von AddA und RecB in *B. subtilis* bzw. *E. coli* hat zudem einen stark negativen Einfluss auf die Häufigkeit der HR-Frequenz (Zuñiga-Castillo *et al*., 2004). In dem Alphaproteobakterium *Rhizobium etli* scheint die HR-vermittelte Integration von Plasmid-DNA hingegen insbesondere von RecF abhängig zu sein (Zuñiga-Castillo et al., 2004). Eine transiente Überexpression nativer Gene des Rekombinationsapparates wie *recA*, *addAB* und *recF* ist daher eine naheliegende Möglichkeit, die HR-Frequenz in *S. melilioti* zu steigern. Alternativ wäre zudem eine Inaktivierung von ssDNA-Exonukleasen denkbar, die sich in *E. coli* negativ auf die Häufigkeit von HR-Ereignissen auswirken (Feschenko *et al*., 2003). In Anbetracht der hohen Effizienz CRISPR/Cas9-induzierter HR in Proteobakterien (Jiang *et al*., 2015; Pyne *et al*., 2015), Firmicuten (Altenbuchner, 2016; Zhang *et al*., 2016; Westbrook *et al*., 2016; Wang *et al*., 2015; Wang *et al*., 2016b; Oh und van Pijkeren, 2014) und Actinobakterien (Cobb und Wang, 2015; Wang *et al*., 2016a) stellt jedoch das CRISPR-System die vielversprechendste Technik dar, um *Genome Editing* auch in Alphaproteobakterien zu erleichtern. In den meisten Prokaryoten werden Cas9-vermittelte DNA-Strangbrüche via HR-basierter Mechanismen repariert (Bowater und Doherty, 2006; Shuman und Glickman, 2007), und CRISPR-basiertes *Genome Editing* macht sich diese Eigenschaft zunutze (Kap. 2.1.1.1). Die Etablierung des CRISPR/Cas9-Systems in *S. meliloti* bietet sich insbesondere zur Optimierung der CFGE-Methode an. Die potentiell sehr hohe Effizienz des CRISPR-Systems sowie die hohe Letalität bei *self-targeting* unmodifizierter Ziel-Sequenzen stellen hierbei zudem Marker-freie Selektionstrategien in Aussicht (Kap. 2.1.1.1). NHEJ birgt in *S. meliloti* weiterhin das Potential, Cas9-vermittelte DSBs mit Hilfe linearer Donor-DNA zu reparieren (Kobayashi *et al*., 2008). In einem NHEJ/CRISPR-basierten *Editing*-Ansatz wäre folglich nur noch die Herstellung und Elektroporation eines linearen DNA-Fragments erforderlich.

3.1.2 Cre/*lox*-Rekombinationssysteme ermöglichen substantielle Genom-Umstrukturierungen

Im Rahmen dieser Arbeit wurden Cre/*lox* basierte Rekombinationssysteme zur Deletion, Inversion und Fusion großer DNA-Moleküle in *S. meliloti* entwickelt (Abb. 3.1). Basierend auf einem von Harrison *et al.* (2011) bereitgestellten Rekombinase-Expressionssystem wurden die *S. meliloti*-Stämme SmCre und SmCreΔhsdR konstruiert, die eine stabile Integration des Rekombinase-Gens unter Kontrolle des Taurin-induzierbaren Promotors P*tauXY* aufweisen (Kap. 6.3). Nicht-induzierte Zellkulturen weisen aufgrund der geringen, basalen Expressionsstärke der nativen Promotorregion keine signifikante Rekombinaseaktivität auf (Harrison *et al*., 2011). Das System zeichnet sich daher als robustes *Engineering*-Werkzeug für komplexe Rekombinationsstrategien aus.

Cre ermöglicht unabhängig spezieller Wirtsfaktoren die ortsspezifische Rekombination zur Deletion „gefloxter“ DNA-Fragmente (Kap. 2.1.2.2). Die in dieser Arbeit hergestellten Basis-Konstrukte pCreDelL und pCreDelR sind universell nutzbar und Teil eines standardisierten Deletionssystems für *S. meliloti*. Dieses lässt sich mit geringem Klonierungsaufwand auf beliebige Genomabschnitte adaptieren. Seine Funktion wurde in dieser Arbeit anhand der Exzision eines ~24 kb großen Gen-Clusters demonstriert. Um die das DNA-Fragment flankierenden Sequenzbereiche mittels einfacher HR sukzessive zu präparieren, wurde die Cre-Erkennungssequenz *loxR* mit mutiertem RE-Element via CFGE stromabwärts der Region integriert (Abb. 3.1A). Dies ermöglichte eine unmittelbare Insertion von pCreDelL (*loxL*) stromaufwärts des Gen-Clusters, ohne eine Fehlintegration über Plasmid-stämmige Sequenzen zu provozieren. Aufgrund der parallelen Anordnung der *lox*-Sequenzen führte die anschließende Cre-Rekombination zur Exzision der „gefloxten“ Region und infolgedessen zu deren Deletion. Die Positivrate lag bei 67%. Die *sacB*-vermittelte Saccharose-Selektion kann somit als zuverlässige Selektionsmethode in ortsspezifischen Rekombinasesystemen angesehen werden.

Hauptmerkmale des etablierten Cre/*lox*-Deletionssystems sind somit ein stark reduzierter Klonierungsaufwand, eine hohe Positivrate sowie eine minimale Deletionsregion, die lediglich eine *loxLR*-Sequenz mit reduziertem Rekombinationspotential aufweist und daher *lox*-vermittelte Folgereaktionen ermöglicht. Ein Flp/*FRT*-basiertes System, mithilfe dessen großangelegte Deletionsstudien in *S. meliloti* durchgeführt wurden, limitiert hingegen aufgrund im Genom verbleibender Selektionsmarker eine sequentielle Anwendung (Milunovic *et al*., 2014). Auch ein für *A. tumefaciens* entwickeltes Cre/*loxP*-Deletionssystem bedarf einer vergleichsweise aufwändigen Klonierung (Liu *et al*., 2016). Eine *loxP* Sequenz, die im Genom verbleibt, verwehrt zudem ortsspezifische Folgereaktionen. Das in dieser Arbeit etablierte Cre/*lox*-Deletionssystem hat somit grundlegende Vorzüge gegenüber bekannten Rekombinasesystemen in α-Rhizobien. Großes Potential liegt jedoch weiterhin in der Reduzierung des Klonierungsaufwandes, der nicht nur mittels einer Optimierung der CFGE-Methode erzielt werden kann. Eine alternative Rekombinationsstrategie, die überhaupt keiner Klonierung bedarf, wurde zuletzt mittels eines modifizierten Gin-Rekombinasesystems zur Deletion in humanen Zellen etabliert (Chaikind *et al*., 2016). So ermöglichte das Fusionskonstrukt RecCas9, ein dCas9-Gin-Hybrid, die sgRNA-vermittelte Deletion eines chromosomalen Fragments ohne vorherige Insertion der Erkennungssequenzen. Dies ist möglich, da die Serin-Rekombinase mit Unterstützung von dCas9 auch 20 bp kleine pseudo-Erkennungssequenzen als Substrat akzeptiert (Chaikind *et al*., 2016). Die Fusion von dCas9 mit ortsspezifischen Rekombinasen könnte auch *Engineering*-Verfahren in Prokaryoten erheblich erleichtern.

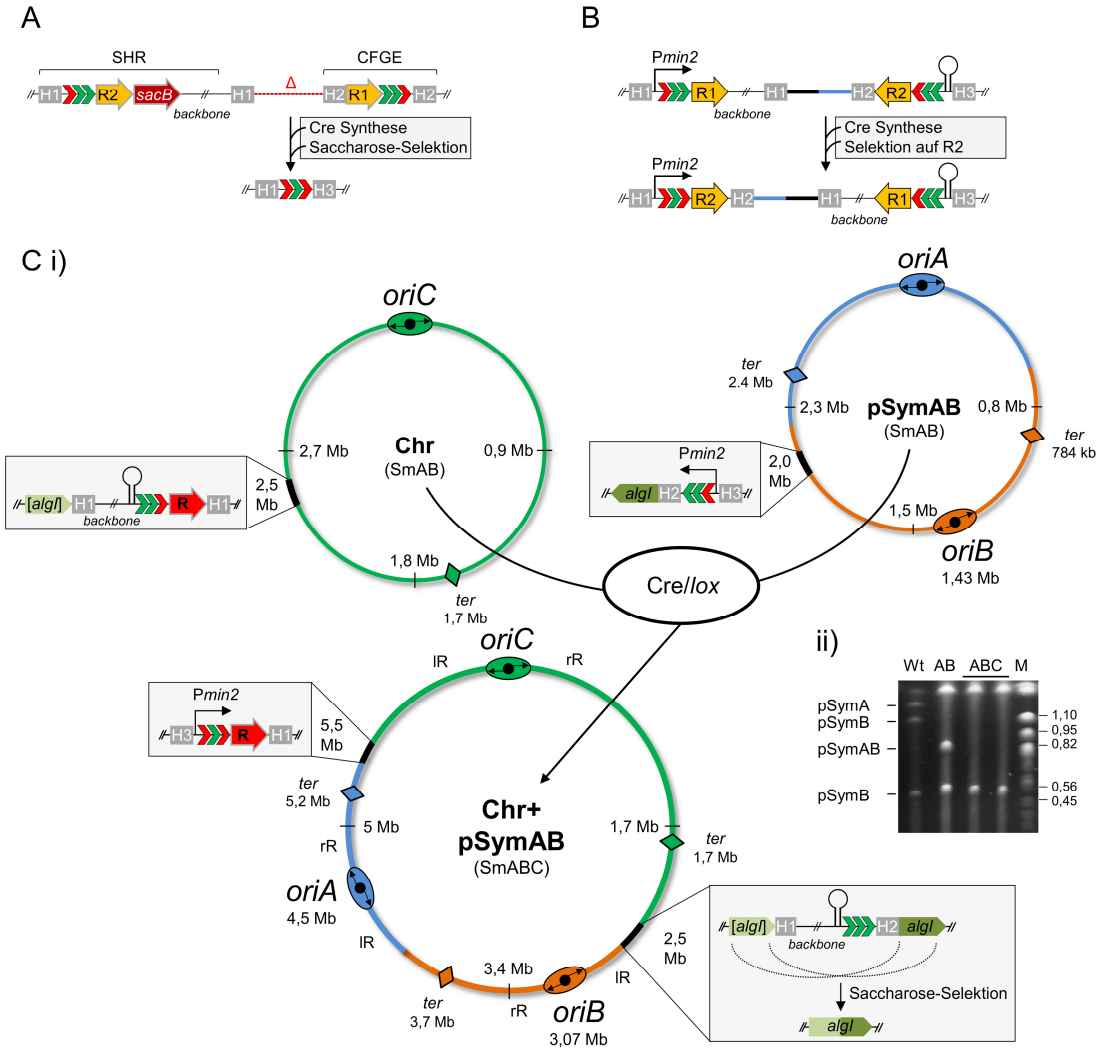


Abbildung 3.1: Schematische Darstellung entwickelter Cre/*lox*-Rekombinationssysteme für *Genetic Engineering* in *S. meliloti*. Dreigliedrige Pfeile kennzeichnen die Orientierung von *lox*-Sequenzen mit Mutationen (rot) im linken (*loxL*) oder rechten Arm (*loxR*). H1-3: Homologe Regionen. Ringstruktur: Transkriptionsterminator. **A:** Cre/*lox*-vermittelte Deletion großer DNA Fragmente. In einem ersten Schritt wird *loxR* mittels CFGE stromabwärts der zu deletierenden Region (Δ) integriert. Eine einfache HR führt zur anschließenden Integration von *loxL*. Der Vektor umfasst außerdem das Levansucrase-kodierende Gen *sacB* und ermöglicht somit die Saccharose-Selektion auf den Verlust von Region Δ nach Cre-Induktion. R1/2: Resistenzkassetten. **B:** Cre/*lox*-vermittelte Inversion großer DNA-Fragmente. Das im Rahmen der Masterarbeit von Meike Brennecke experimentell überprüfte Inversionssystem basiert auf Cre-Erkennungssequenzen in Kopf-zu-Kopf-Orientierung und einem Wechsel der Antibiotika-Resistenz nach erfolgreicher Inversion der „gefloxten" Region. R1/2: Resistenzgene. **C:** Cre/*lox*-Fusion der bereits vereinten Megaplasmide (pSymAB) mit dem Chromosom in *S. meliloti* SmAB. Eine detaillierte Beschreibung der Konstruktion ist in Kapitel 6.1 zu finden. i) Ein Transkriptionsterminator verhindert die vorzeitige Expression eines Resistenzgens (R) auf pSymAB. Eine *sacB*-basierte Gegenselektion ermöglicht nach Cre-Rekombination die HR-vermittelte Deletion einer funktionellen *loxP*-Sequenz, des Transkriptionsterminators, und des Plasmid-B*ackbones*. Der resultierende Stamm SmABC weist Resistenzkassetten auf, die lediglich inaktivierte *loxLR*-Sequenzen enthalten. Die Replichor-Verhältnisse (lR:rR) von Chr+pSymAB sind nahezu ausgeglichen. *oriA*/*B*/*C*: Replikationsursprung von pSymA, pSymB und Chromosom. *ter*: Terminus-Region. Chr: Chromosom. ii) Pulsfeld-Gelelektrophorese zum Nachweis der Replikon-Fusionen. Die Linearisierung genomischer DNA aus *S. meliloti* Rm1021 (Wt), SmAB (AB) und SmABC (ABC) mit dem Restriktionsenzym PacI resultiert in einem spezifischen Bandenmuster. Die Auftrennung erfolgte im Bereich 0,5 – 1,5 Mb. Wt: 1,3 Mb (pSymA); 1,15 Mb (pSymB Fragment); 0,53 Mb (pSymB Fragment). SmAB: 0,84 Mb (pSymAB Fragment); 0,53 Mb (pSymB Fragment). SmABC: 0,53 Mb (pSymB Fragment). M: Yeast Chromosome PFG Marker (NEB).

Mittels spezifischer Mutationen im *Spacer*-Bereich von *loxP* kann das Repertoire alternativer *lox*-Sequenzen erweitert werden (Langer *et al*., 2002; Lee und Saito, 1998; Missirlis *et al*., 2006). In dieser Arbeit konnten vier *Spacer*-Mutanten identifiziert werden, die in *S. meliloti* exklusiv mit sich selbst und in hoher Effizienz rekombinieren (Kap. 6.3). Die Ergebnisse widersprechen dem Konsens, dass heterospezifische *lox*-Sequenzen aufgrund der relativ kleinen *Crossover*-Region zu einer signifikant reduzierten Reaktionsspezifität bei verminderter Rekombinationsfrequenz neigen (Turan *et al*., 2011). Die Möglichkeit, mittels mutierter RE/LE-Elemente die Direktionalität der Cre-Reaktion zu forcieren, stellt zudem einen großen Vorteil gegenüber dem Flp/*FRT*-Rekombinasesystem dar und erlaubt die Entwicklung komplexer Rekombinationsstrategien. Da die Frequenz einer *loxLR*-vermittelten Rekombination gegenüber *loxP* jedoch nur ~50-fach reduziert ist, bedürfen Rekombinationsprodukte entsprechender SmCre-Derivate jedoch einer stetigen Selektion. Um unbeabsichtigte Rekombinationsreaktionen zu unterbinden, sollte daher langfristig eine Kombination orthogonaler und unidirektionaler Rekombinasesysteme in Betracht gezogen werden. Insbesondere Integrase-Systeme wie phiC31 und IntA, die bereits in *S. meliloti* bzw. *R. etli* etabliert wurden (diCenzo *et al*., 2013; Hernández-Tamayo *et al*., 2016), bieten sich für entsprechende Integrations-Anwendungen an.

In noch unpublizierten Arbeiten wurden Systeme zur Inversion und Replikon-Fusion entwickelt, die das Repertoire Cre/*lox*-basierter Rekombinationsstrategien in *S. meliloti* vervollständigen. Im Falle einer Anordnung der *loxP*-Sequenzen in „Kopf-zu-Kopf"-Orientierung vermittelt Cre eine Inversion der flankierten Sequenz (Kap. 2.1.2.2). Ein universell anwendbares, Cre/*lox*-basiertes Inversionssystem wurde konzipiert und im Rahmen der Masterarbeit von Meike Brennecke erfolgreich anhand der Inversion einer 24 kb großen genomischen Region demonstriert (Abb. 3.1B). Die serielle Anwendung dieses Systems ist jedoch problematisch, da zuvor integrierte Elemente und eine native *loxP*-Sequenz prinzipbedingt im Genom verbleiben.

Obwohl Cre die Exzisions- gegenüber der Integrationsreaktion bevorzugt (Turan *et al*., 2013), konnten im Rahmen dieser Arbeit Replikon-Fusionsreaktionen zwischen 1,3 - 3,7 Mb großen Replikons in hoher Effizienz erzielt werden. Nach dem Vorbild eines *S. meliloti*-Derivats, das homologe Co-Integrationen zwischen den duplizierten Genregionen von *nodPQ* (pSymA und pSymB) und *algI* (pSymB und Chromosom) aufweist (Guo *et al*., 2003), wurde mit Hilfe einer Cre/*lox*-basierten Klonierungsstrategie die vollständige Fusion des Genoms von *S. meliloti* erzielt. Hierzu wurden zunächst die Megaplasmide pSymA und pSymB (SmAB), und daraufhin das Chromosom mit pSymAB ineinander integriert, woraus schließlich das 6,68 Mb große Replikon des *S. meliloti*-Stammes SmABC hervorging (Abb. 3.1C). (Sekundär-) Chromosome weisen eine evolutionär geprägte Symmetrie auf (Lobry und Louarn, 2003; Hendrickson und Lawrence, 2006; Touzain *et al*., 2011), die sich in *S. meliloti* in einer gleichmäßigen Anordnung von Replikationsursprung und Terminus-Region sowie der spezifischen Verteilung weiterer

Symmetrie-Elemente äußert (Hendrickson und Lawrence, 2006). Daher war es fraglich, ob *S. meliloti* genomische Modifizierungen in einem Maßstab toleriert, der diese Balance erheblich stören würde. Die erfolgreiche Herstellung von sechszehn *S. meliloti* SmAB-basierten Replikon-Fusionsstämmen, die im Rahmen der Masterarbeit von Meike Brennecke durchgeführt wurde, stellt daher ein Schlüsselexperiment dar. *S. meliloti* SmABC weist ausgeglichene Replichor-Verhältnisse und in Vollmedium eine mit dem Wildtyp vergleichbare Fitness auf (Brenecke, 2016). Neue SmABC-Derivate mit variablen Replichor-Verhältnissen deuten hingegen eine Korrelation zwischen Replikon-Symmetrie und Wachstumsrate an, wobei die Wachstumsrate bei zunehmend unausgeglichenen Replichor-Verhältnissen abnimmt (Brennecke, 2016). Dennoch gelang die Herstellung fusionierter Replikons mit stark unausgeglichenen Replichor-Verhältnissen, und selbst *S. meliloti*-Derivate, deren Genom eine Unterbrechung der *oriV-ter*-Abfolge aufweist, sind vital (Brennecke, 2016).

S. meliloti-Stämme mit gänzlich neuer Genomarchitektur sind lebensfähig und beweisen einen in Hinblick auf die Replikon-Organisation und -Struktur erstaunlich hohen Grad an Flexibilität. Dieser gewährt großen Spielraum in zukünftigen *Engineering*-Anwendungen, ermöglicht aber auch neue Ansätze zur funktionellen Untersuchung der Chromosomenbiologie. Ein umfassendes Verständnis der zugrunde liegenden Mechanismen ist notwendig, um die ambitionierten Ziele der Synthetischen Biologie in Hinblick auf die Konstruktion subchromosomaler DNA-Moleküle und artifizieller Chromosome zu verwirklichen (Kap. 2.1.2). Die zwei- und dreidimensionale Genomorganisation bakterieller Modellorganismen wie *E. coli, B. subtilis* und *Vibrio cholerae* konnte bereits in hohem Detail aufgeklärt werden (Wang und Rudner, 2014; Dekker *et al*., 2013). So bestimmen Makrodomänen, wie sie in *E. coli* beschrieben sind (Dorman, 2013), die Topologie des Genoms und haben regulatorische Auswirkungen bis auf die Ebene der Genexpression (Slager und Veening, 2016). Genetische Werkzeuge zur DNA-Modifikation auf makromolekularer Ebene sind notwendig, um solche übergeordneten Strukturen mit spezifischen Funktionen in Beziehung zu bringen. In *V. cholerae*, einem Modellorganismus zur Untersuchung eines mehrteiligen Genoms, ermöglichten erst Rekombinase-vermittelte Inversionen großer chromosomaler Regionen die Identifizierung einer Checkpoint-Kontrolle, die die Replikationsinitiation des Sekundärchromosoms reguliert (Val *et al*., 2016). Vergleichbare Erkenntnisse zur Organisation des mehrteiligen Genoms rhizobieller Spezies beschränken sich auf die Genomarchitektur und sind unzureichend (Guo *et al*., 2003; Mavingui *et al*., 2002). Im Rahmen dieser Arbeit entwickelte *Engineering*-Werkzeuge und hergestellte Fusionsstämme unterstützen zukünftige Projekte, in welchen mittels fluoreszenzmikroskopischer Verfahren zur Replikon-Markierung (Lau *et al*., 2003) sowie mit Hilfe einer Hi-C-basierten Methode zur topologischen Genomanalyse (Dekker *et al*., 2013) die Aufklärung der raumzeitlichen und strukturellen Genomorganisation angestrebt wird. Dies ermöglicht die Ableitung regulatorischer Prinzipien und fundamentaler Design-Regeln eines mehrteiligen Genoms.

3.2 Entwicklung neuer *Genetic Engineering*-Ansätze mit hohem Anwendungspotential in biotechnologischen und biologischen Fragestellungen

3.2.1 *repABC*-basierte Replikationsursprünge vermitteln Chromosomen-ähnlichen Eigenschaften

Insbesondere schnell wachsende α-Rhizobien besitzen extra-chromosomale Replikons, die in der Regel der RepABC-Familie angehören und im Unterschied zu typischen Plasmiden eine sehr geringe Kopienzahl aufweisen (Kap. 2.2.1). Am Beispiel von *S. meliloti* konnte erstmals die Segregations-Dynamik eines mehrteiligen Genoms aus der Gruppe der Alphaproteobakterien im Detail analysiert werden. *Fluorescent repressor operator system* (FROS)-basierte Lokalisierungsstudien zu pSymA und pSymB offenbarten dabei eine hochgradig regulierte Koordination, die zu einer präzisen Verteilung der duplizierten Replikons auf die Tochterzelle führt und dem Segregationsprozess der Megaplasmide chromosomalen Charakter verleiht (Kap. 6.2). Um herauszufinden, ob das zeitlich und räumlich abgestimmte Verhalten von pSymA exklusiv durch den *repABC*-basierten Replikationsursprung determiniert ist, wurde in dieser Arbeit das artifizielle Mini-Replikon pART auf Basis der zugrundeliegenden *repA2B2C2*-Kassette von pSymA konstruiert. Erwartungsgemäß verhinderten Inkompatibilitätsmechanismen des Replikationsursprungs eine Transformation des 11,3 kb großen Replikons in *S. meliloti* Rm1021, wohingegen seine autonome Replikation im pSymA-freien *S. meliloti*-Stamm Sma818 nachgewiesen werden konnte. Tatsächlich imitiert pART das 1,3 Mb große Megaplasmid sowohl hinsichtlich der einfachen Kopienzahl, als auch der raumzeitlichen Koordination während des Zellzyklus. Lediglich die Stabilität der subpolaren Lokalisierung nach erfolgter Segregation erscheint weniger stabil als bei pSymA, was der geringeren Trägheit des 120-fach kleineren DNA-Moleküls geschuldet sein könnte.

Demzufolge stellt die lediglich ~4 kb umfassende *repA2B2C2*-Region einen vollwertigen Replikationsursprung dar, der die Chromosomen-ähnlichen Eigenschaften eines nativen Megaplasmids auf einen nicht-replikativen Vektor überträgt. Dies warf die Frage auf, ob diese Eigenschaften auch durch *repABC*-Kassetten artfremder Rhizobien vermittelt wird. Infolgedessen wurden in dieser Arbeit zwölf *repABC*-Regionen der α-Rhizobien *R. etli, Mesorhizobium loti, A. tumefaciens, Sinorhizobium medicae* und *Sinorhizobium fredii* auf die Eigenschaft hin überprüft, einem nicht-replikativen Vektor in *S. meliloti* die Fähigkeit zur autonomen Replikation zu verleihen. Während dies im Falle der *S. medicae*- und *S. fredii*-stämmigen *repABC*-Kassetten vermutlich aufgrund ihrer nahen Verwandtschaft zu *S. meliloti* nicht gelang, konnten *R. etli*, *M. loti* und *A. tumefaciens* als vielversprechende Spender *S. meliloti*-kompatibler Replikationsursprünge identifiziert werden. Diese Beobachtung war wegbereitend für die Entwicklung einer *repABC*-basierten Vektor-Familie sowie einer neuen *in-vivo* Klonierungsmethode.

3.2.1.1 pABC Shuttle-Vektoren sind artifizielle Mini-Replikons mit modularem Aufbau

Die in Anlehnung an den Replikationsursprung rhizobieller Megaplasmide benannten *Shuttle*-Vektoren setzen sich aus bis zu fünf Modulen zusammen (Kap. 6.4 und Abb. 3.2). Das oriVSm Modul beinhaltet *repABC*-Kassetten aus *R. etli*, *M. loti* oder *A. tumefaciens*, die in *S. meliloti* als Replikationsursprung fungieren. Um die transkriptionelle Regulation des *repABC*-Operons zu gewährleisten, wird seine Promotor-Region von einer synthetischen Terminator-Region flankiert (synTer-MCS). Diese enthält vier Kopien der *E. coli*-stämmigen Transkriptionsterminatoren *rrnBT1* und *rrnBT2* (Orosz *et al.*, 1991), sowie eine MCS mit mehreren Restriktionsschnittstellen geeigneter Seltenschneider. Das nachfolgende oriVEc-Modul stellt *E. coli*-kompatible Replikationsursprünge bereit, die eine *E. coli*-basierte Konstruktion sowie konventionelle Klonierungstechniken zum Beladen der pABCs ermöglichen. Den Transfer eines pABCs mittels des RK2/RP4-Konjugationssystems nach *S. meliloti* erlaubt das optionale oriT-Modul, wohingegen das AR-Modul standardisierte Antibiotikaresistenzkassetten zur Selektion in *E. coli* und *S. meliloti* bereitstellt. Insgesamt wurden 24 Komponenten hergestellt und in eine DNA-Bibliothek aufgenommen. Darin enthaltene „Bausteine“ sind von Modul-spezifischen Linker-Sequenzen flankiert, die eine LCR-vermittelte Assemblierung der Modul-Fragmente mit Hilfe definierter Oligonukleotid-Sets ermöglichen. Bis zu 12 kb große pABC-Derivate wurden assembliert und in *E. coli* DH5α transformiert. Positivraten von >75% bestätigen die hohe Assemblierungseffizienz des optimierten LCR-Protokolls (de Kok *et al.*, 2014) sowie die Eignung der eingesetzten Linker-Sequenzen. Das entworfene Assemblierungssystem erfüllt somit die Ansprüche moderner Vektor-Plattformen (Durante-Rodríguez *et al.*, 2014) und erleichert die Integration alternativer bzw. neuer Modul-Typen.

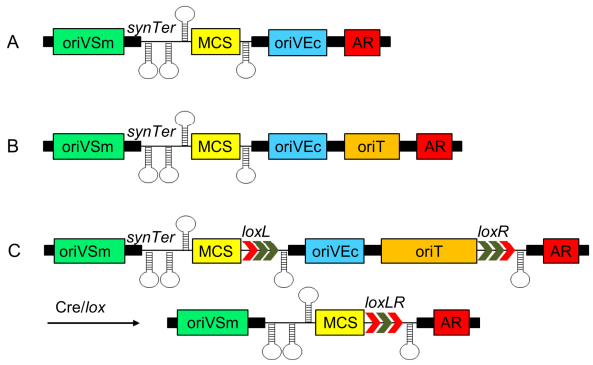


Abbildung 3.2: In dieser Arbeit hergestellte pABC Vektoren setzen sich aus bis zu fünf Modulen zusammen, die durch spezifische Linker (schwarz) miteinander verbunden sind. Während pABCs des Typs A in *S. meliloti* elektroporiert werden müssen, erlauben Vektoren des Typs B und C aufgrund des oriT-Moduls den konjugalen Transfer zwischen *E. coli* und *S. meliloti*. Typ C ermöglicht die Cre/*lox*-vermittelte Deletion der Module oriVEc und oriT, und vermeidet potentiell problematische Sequenz-Homologien in multi-pABC-Systemen. Ring-Struktur: Transkriptionsterminator des synTer-MCS-Moduls.

Sechszehn individuelle pABC *Shuttle*-Vektoren wurden in Zusammenarbeit mit Marcel Wagner auf Basis acht verschiedener *repABC*-Kassetten hergestellt und begründen eine neue Familie replizierender *Shuttle*-Vektoren für *S. meliloti* Rm1021 (Tab. 3.1).

Tabelle 3.1: Charakterisierte pABC *Shuttle*-Vektoren. Im Vergleich zu ihren Vorgänger-Derivaten pABCa-c weisen pABC1-6 Modul-spezifische Linker-Sequenzen zur standardisierten Assemblierung auf. pABC1-lox-Derivate unterscheiden sich durch ihr oriT-Modul, das im Falle von pABC1-loxB eine zusätzliche Kopie des Transkriptionsterminators *rrnBT2* enthält (vgl. Abb. 3.2C).

		pABC Modul-Komponenten						
	Name	oriVSm (Quelle)	synTer-MCS (*synTer* Fragment)	oriVEc (Quelle)	oriT	AR	Größe (kb)	Design Typ
Vorgänger	**pABCa**	pMlb (*M. loti*)	*synTer1*	p15A	-	Gm	6,7	A
	pABCb	p42b (*R. etli*)	*synTer2*	pMB1	-	Sp	6,5	A
	pABCc	p42d (*R. etli*)	*synTer3*	pSC101*	-	Km	11,5	A
	pABCc-mob	p42d (*R. etli*)	*synTer3*	pSC101*	✓	Km	12,0	B
Assemblierung aus Modul-Bibliothek	**pABC1**	p42b (*R. etli*)	*synTer1*	pMB1	-	Sp	6,7	A
	pABC1-loxA	p42b (*R. etli*)	*synTer1-lox*	pMB1	✓	Sp	9,5	C
	pABC1-loxB	p42b (*R. etli*)	*synTer1-lox*	pMB1	✓	Sp	9,6	C
	pABC1-sT2	p42b (*R. etli*)	*synTer2*	pMB1	-	Sp	6,7	A
	pABC1-sT3	p42b (*R. etli*)	*synTer3*	pMB1	-	Sp	6,7	A
	pABC2	pMla (*M. loti*)	*synTer2*	p15A	-	Gm	7,2	A
	pABC3	pTi (*A. tumefaciens*)	*synTer3*	pSC101*	-	Hyg	8,6	A
	pABC4	p42e (*R. etli*)	*synTer1*	pMB1	-	Sp	7,0	A
	pABC5	p42f1 (*R. etli*)	*synTer2*	pMB1	-	Sp	7,2	A
	pABC6	p42f2 (*R. etli*)	*synTer2*	pMB1	-	Sp	7,3	A

p42f1/2: *repA1B1C1*- bzw. *repA2B2C2*-Region von p42f. pSC101*: pSC101-Derivat.

Es wurden vier Replikationsursprünge basierend auf p42b-, p42d-, p42e-, und pMlb-stämmigen *repABC*-Kassetten identifiziert (korrespondierend mit pABCa-c, pABC1 und pABC4), die eine außergewöhnlich hohe Vererbungsstabilität sowie einfache Kopienzahl in *S. meliloti* Rm1021 vermitteln.

Letzteres wurde in fluoreszenzmikroskopischen Untersuchungen nachgewiesen. Hierzu wurde ein auf der *tetR-YFP-tetO*- (Lau *et al*., 2003) und *parB-mChr-parS*- (Schwartz und Shapiro, 2011) Interaktion beruhendes System zur *in-vivo* Replikon-Markierung etabliert. Der resultierende *S. meliloti*-Stamm JDSm106 ermöglicht eine Taurin-induzierbare Expression der fluoreszenten Reporterproteine und somit eine simultane Markierung zwei individueller Replikons. JDSm106-Derivate, die mit *parS*- bzw. *tetO*-Sequenzen markierte pABCa-c- bzw. pABC1- und pABC4-Derivate propagieren, weisen ein bis zwei individuelle Fokusse auf. Dies lässt auf eine einfache Kopienzahl der ggf. replizierten Mini-Replikons schließen, die außerdem mit Hilfe der qPCR-basierten ΔΔCt-Methode (Livak und Schmittgen, 2001) anhand korrespondierender oriVSm Modul-Konstrukte verifiziert wurde. Dieser Befund ist insofern überraschend, als dass die in der Literatur als *single*- bis *low-copy* beschriebene (Pinto *et al*., 2012; Cervantes-Rivera *et al*., 2011; Cevallos *et al*., 2008; McAnulla *et al*., 2007; Pappas und Winans, 2003a), unpräzise Angabe zur Kopienzahl der korrespondierenden Megaplasmide durchaus Abweichungen erwarten ließ.

Die Vererbung der pABC-*eGFP*-Derivate wurde hingegen mittels Fluoreszenz-Durchflusszytometrie überprüft. Hierzu wurden entsprechende *S. meliloti*-Kulturen über 92 Stunden in nicht-selektivem TY-Medium inkubiert und regelmäßig verdünnt. Da ein positives EGFP-Signal mit einer erfolgreichen Vererbung des Vektors korreliert, erlaubt die Methode eine Quantifizierung des Anteils positiver Zellen. Derivate von pABCa-c, pABC1 und pABC4 konnten nach 92-stündigem Wachstum in über 97% der analysierten Zellen nachgewiesen werden, während der klassische *low-copy* Vektor pPHU231 in 45% der Zellkultur verloren ging. Diese Ergebnisse stehen mit einem Selektionsmarker-basierten Stabilitätstest im Einklang, der die stabile Vererbung korrespondierender oriVSm Modul-Vektoren validiert. Einen Sonderfall hinsichtlich der Vererbungsstabilität stellt pABC3 dar. Unabhängige *S. meliloti*-Kulturen, die das *repABC*$_{pTi}$-basierte Derivat propagieren, wiesen bereits nach 48-stündiger Inkubation Verlustraten in Höhe von ~100, ~50 und ~0% auf. Für die unstetige Vererbung könnte ein *Quorum-Sensing* (QS)-abhängiger Mechanismus verantwortlich sein. In *A. tumefaciens* verstärkt der LuxR-ähnliche QS-Regulator TraR als Transkriptionsfaktor die Expression des *repABC*-Operons und interagiert wahrscheinlich mit RepC, was eine Modulation der Kopienzahl des Megaplasmids pTi zur Folge hat (Pappas und Winans, 2003a). Da die pABC3 zugrundeliegende *repABC*$_{pTi}$-Region alle bekannten *tra*-Boxen umfasst und auch *S. meliloti* mindestens acht LuxR-ähnliche Proteine aufweist (Charoenpanich *et al*., 2013), die als potentielle Bindeproteine in Frage kommen, erscheint eine Dysregulation der Kopienzahlkontrolle als Folge einer unspezifischen Interaktion dieser Elemente als plausible Ursache für das heterogene Verhalten des *Shuttle*-Vektors. Bei der Suche nach alternativen *repABC*-Kassetten aus weiteren Rhizobien werden die Replikationsursprünge QS-regulierter Megaplasmide folglich als ungeeignete Kandidaten eingestuft.

Da sowohl die hohe Vererbungsstabilität, als auch die einfache Kopienzahl genannter pABCs und korrespondierender oriVSm Modul-Konstrukte übereinstimmend in komplementären Experimenten nachgewiesen werden konnten, dürfen die erhobenen Daten als äußerst robust und beschriebene Charakteristika als *repABC*-spezifisch bzw. Plasmid-*Backbone*-unabhängig betrachtet werden. Das Repertoire replizierender Vektoren für *S. meliloti*, das sich mit pBBR1- und pRK2-Derivaten bislang auf Plasmide mit mittlerer und geringer Kopienzahl beschränkte (Khan *et al*., 2008; Hübner *et al*., 1991), konnte in dieser Arbeit somit um vier *single-copy* Derivate erweitert werden. Die weite Verbreitung *repABC*-basierter Replikons innerhalb der Rhizobien suggeriert außerdem ein Organismen-übergreifendes Anwendungspotential der pABC Vektoren, das jedoch weitere Kompatibilitäts-Untersuchungen erfordert.

3.2.1.2 Ein konserviertes Motiv im *repABC*-Operon

Fluoreszenzmikroskopisch analysierte pABC-*tetO*/*parS*-Derivate wiesen in unterschiedlichen Phasen des Zellzyklus ein Lokalisierungsmuster auf, das an die hochregulierte Koordination des artifiziellen Mini-Replikons pART erinnert (Kap. 6.4). Diesem Verhalten liegt wahrscheinlich eine fein-justierte Regulation der *repABC*-Expression zugrunde, die mittels eines noch unbekannten Regulators mit dem allgemeinen Zellzyklus synchronisiert sein könnte. Ein Transkriptionsregulator etwa könnte eine konservierte Sequenz im Promotor-Bereich des *repABC*-Operons binden. Tatsächlich konnte in dieser Arbeit mithilfe des Analyse-Tools Glam2 (Frith *et al*., 2008) ein sieben Basenpaare großes Motiv (GAATCGG) in 18 *repABC*-Operons sechs verschiedener Rhizobien identifiziert werden (Abb. 3.3A). Überraschenderweise tritt die Sequenz sowohl einfach im Promotorbereich von *repA*, als auch zweifach in der intergenischen Region zwischen *repB* und *repC* auf. In Sequenzhomologie-Analysen der ctRNA-Region wurde zudem ein 49-60 bp großer Abschnitt mit hoher Sequenzidentität identifiziert, der das repetierende Motiv enthält, dieses aber nicht als solches charakterisiert und in Zusammenhang mit einer Expressions-regulatorischen Funktion bringt (Chain *et al*., 2000). Letzteres liegt jedoch nahe, da die Sequenz eine potentielle Methylierungs-Stelle (GAATC) der DNA-Methylase CcrM enthält, die nicht nur in *C. crescentus*, sondern auch in vielen Rhizobien als integraler Bestandteil der Zellzyklus-Regulation bzw. der DNA-Replikationsinitiation angesehen wird (Wright *et al*., 1997; Kahng und Shapiro, 2001). Im Falle der Megaplasmide pSymA (MacLellan *et al*., 2006), pSymB (Schlüter *et al*., 2013) und p42d (Ramírez-Romero *et al*., 2001) ist außerdem der Transkriptionsstart (TSS) des Operons bekannt und nahezu deckungsgleich mit besagter Sequenz. Obgleich für diese Korrelation ein experimenteller Nachweis aussteht, wurde das Motiv zur Vorhersage des TSS im *repABC*-Operon von pMlb herangezogen. Mit der Intention, die sensible Regulation des Replikationsursprungs zu stören, wurde im Rahmen der Masterarbeit von Carina Happel und in Anlehnung an die regulatorische Funktionsweise des *lac*-Operons (Oehler *et al*., 1994) eine *lacO1*-Box unmittelbar stromabwärts des vorhergesagten TSS inseriert (Abb. 3.3B i). Das resultierende Plasmid pMlb_lacO mit *repABC*$_{pMlb}$-*lacO*-basiertem Replikationsursprung ist in *S. meliloti* replikativ, verliert diese Eigenschaft jedoch bei konstitutiver *lacI*-Expression, und erlangt sie durch Zugabe von IPTG wieder zurück (Kap. 6.4).

Nach der Etablierung thermosensitiver Replikationsinitiatoren (Withers und Bernander, 1998), limitierender Wachstumsbedingungen (Ferullo *et al*., 2009; de Nisco *et al*., 2014) und einer dCas9-vermittelten Replikationsblockade (Wiktor *et al.,* 2016) stellt das in dieser Arbeit entwickelte System einen neuen Ansatz zur kontrollierbaren Replikationsinitiation dar. Ein reprimierbarer *repABC*-Replikationsursprung könnte sich zur Unterbrechung des Zellzyklus und somit zur Synchronisierung von *S. meliloti*-Zellkulturen eignen. In dieser Arbeit wurde das System jedoch verwendet, um die genomische Integration eines induziert-replikativen DNA-Moleküls zu erzielen.

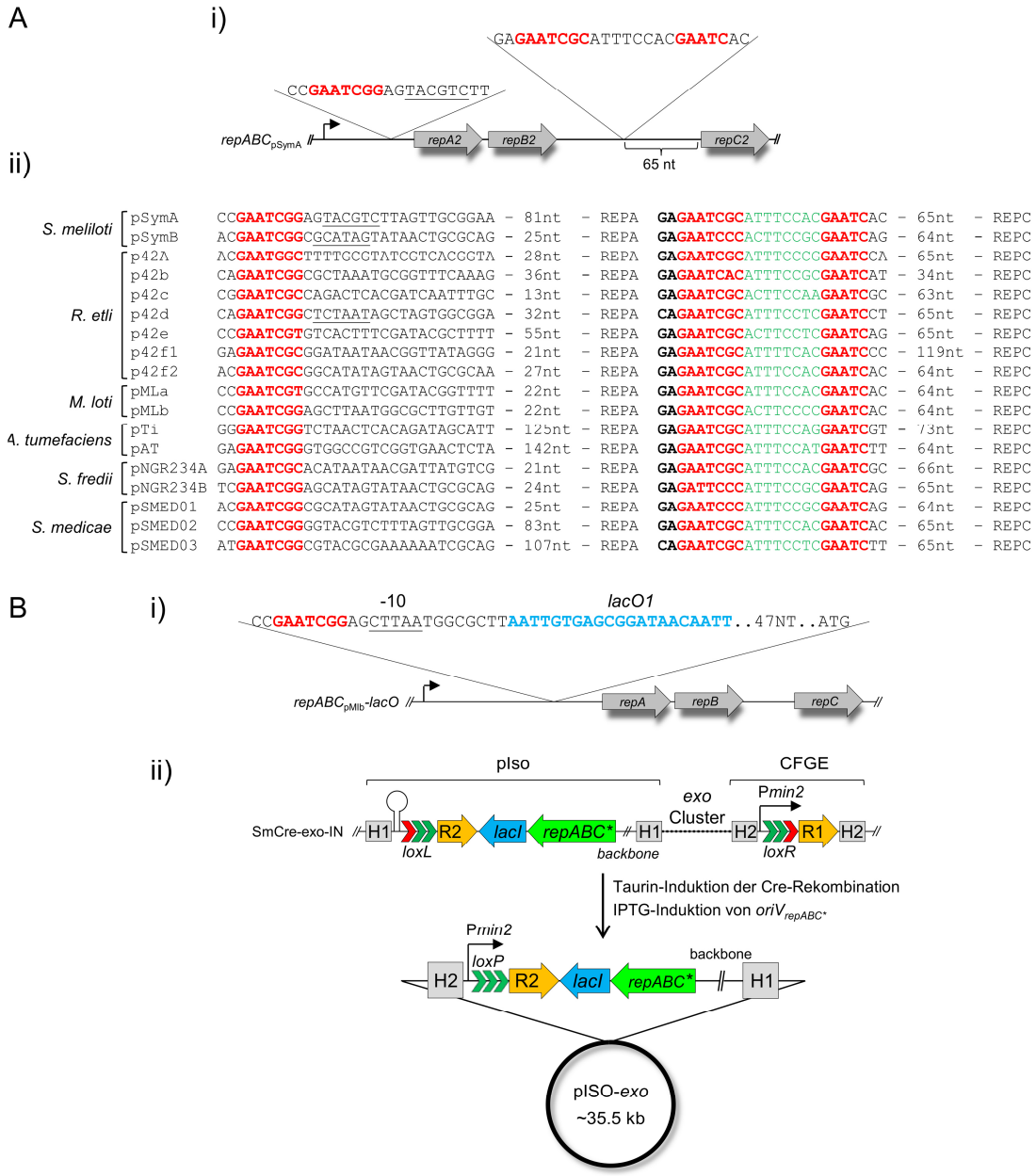


Abbildung 3.3: *In-vivo* Klonierung großer DNA-Fragmente in *S. meliloti*. **A:** *repABC*-Operons verschiedener Rhizobien weisen ein konserviertes Sequenzmotiv auf. i) Am Beispiel des pSymA-stämmigen *repABC*-Operons ist die Position des Motivs (rot) im Promotor-Bereich sowie in der intergenischen Region zwischen *repB* und *repC* veranschaulicht. In Letzterer ist zudem ein um 8 bp erweitertes Motiv erkennbar (grün), das von einer weiteren, unvollständigen GAATC-Sequenz flankiert wird. ii) Das Motiv findet sich in 18 *repABC*-Kassetten sechs rhizobieller Arten. Für pSymA, pSymB und p42d ist zudem die Position der -10 Region des Promotors bekannt (unterstrichen). **B:** ABC *Cloning* ermöglicht die *in-vivo* Klonierung großer DNA-Fragmente. i) Promotor-Region des *repABC*-Operons von pMlb mit integrierter *lacO1*-Sequenz. Der resultierende Replikationsursprung wird durch LacI gehemmt und mit IPTG aktiviert. ii) *S. meliloti*-Stamm SmCre-exo-IN weist eine durch P*min2* exprimierte Gentamicin-Resistenzkassette (R1) sowie eine *loxR*-Sequenz stromabwärts des Gen-Clusters auf. Die Integration von pISO mit $repABC_{pMlb}$-*lacO*-basiertem Replikationsursprung (*repABC**) gelang unter IPTG-freien Bedingungen und führte zur Insertion einer *loxL*-Sequenz stromaufwärts des Gen-Clusters. Die Cre-Induktion führte zur Exzision der „gefloxten" Region. Das resultierende Mini-Replikon pISO-exo besitzt eine konstitutitv exprimierte *lacI*-Kassette und ist daher IPTG-abhängig replikativ. LE/RE-Mutationen der *loxL*/*R*-Sequenzen sind durch dreigliedrige Pfeile mit roter Markierung gekennzeichnet. H1/2: Homologe Regionen zur genomischen Integration. R1/2: Kodierende Sequenzen der Gentamicin-/ Spectinomycin-Resistenzgene.

3.2.1.3 Die Cre/*lox*-vermittelte Relokalisation eines DNA-Fragments ermöglicht die *in-vivo* Klonierung großer Gen-Cluster

Auf Basis des Cre/*lox*-Deletionssystems (Kap. 3.1.2) und mit Hilfe des *repABC*$_{pMlb}$-*lacO*-basierten Replikationsursprungs (Kap. 3.2.1.2) konnte im Rahmen dieser Arbeit eine neue *in-vivo* Klonierungsmethode (ABC *Cloning*) in *S. meliloti* etabliert werden (Kap. 6.4). Zur Demonstration dieser Methode wurde das ~24 kb große *exo*-Gen-Cluster auf pSymB mit parallel angeordneten *loxL*/*R* Sequenzen flankiert (Abb. 3.3B ii). Hierzu wurde die von einer *loxR*-Sequenz unterbrochene, aber funktionale Gentamicin-Resistenzkassette stromabwärts der *exo*-Gene mittels CFGE inseriert. Das Isolationskonstrukt pIso weist eine konstitutive *lacI*-Kassette auf, die die Funktion des modifizierten Replikationsursprungs in IPTG-freiem Medium unterdrückt und eine genomische Integration stromaufwärts der Gen-Region ermöglicht. Weiterhin stellt pIso mit einem Spectinomycin-Resistenzgen, dessen vorzeitige Expression durch einen Transkriptionsterminator unterbunden wird, ein System zur Selektion auf eine erfolgreiche, Cre-vermittelte Exzision des Gen-Clusters bereit. Der komplexe Aufbau verdeutlicht den vielseitigen Nutzen des 72 bp kleinen Promotors P*min2*, der eine adäquate Expression multipler, standardisierter Expressionskassetten in demselben genomischen Kontext gewährleistet, ohne eine homologe Co-Integration zu provozieren.

Tatsächlich führte die Zugabe von IPTG während der Cre-vermittelten Exzision des *exo*-Gen-Clusters zur Retention des 35,5 kb großen Rekombinationsprodukts pIso-exo, das analog zu einem pABC Vektor autonom und mit einfacher Kopienzahl replizierte (Abb. 3.3B ii). Während die *loxL*/*loxR*-basierte Cre-Reaktion im Genom erwartungsgemäß eine inaktive *loxLR*-Sequenz hinterließ, verblieb auf dem Mini-Replikon zudem eine native *loxP*-Sequenz, die sich für weitere Cre-Rekombinationen nutzen ließe. Mittels triparentaler Konjugation konnte pIso-exo außerdem nach *E. coli* sowie *A. tumefaciens* transferiert werden, und auch ein anschließender Transfer von *A. tumefaciens* nach *E. coli* gelang (Happel, 2016). ABC *Cloning* ermöglicht demzufolge eine Relokalisierung großer DNA-Fragmente sowie deren Transfer zwischen *E. coli* und α-rhizobiellen Modellorganismen.

3.2.2 *repABC-Engineering* hat biologisches und synthetisches Anwendungspotential

Neben der makromolekularen Organisation des Genoms von *S. meliloti* (Kap. 3.1.2) sind auch die molekularen Mechanismen der raumzeitlichen Koordination individueller Replikons weitgehend unbekannt. pABC Vektoren eignen sich als markierbare Mini-Replikons (Kap. 3.2.1.1) jedoch hervorragend zur Untersuchung dieses Systems. Tatsächlich vermittelt nicht nur der auf die *repABC*$_{pSymA}$-Kassette reduzierte Replikationsursprung von pART die Kerneigenschaften des nativen Megaplasmids auf ein artifizielles Plasmid. Auch das mittels fluoreszenzmikroskopischer Analysen ermittelte

Lokalisierungsmuster der auf heterologen *repABC*-Regionen basierenden pABCs deutet auf eine mit dem Zellzyklus abgestimmte Koordination hin (Kap. 6.4). *repABC*-basierte Mini-Replikons können somit in beliebigen Konfigurationen sowie in Kombination mit pSymA und pSymB, oder an Stelle der nativen Pendants in Megaplasmid-freien Stämmen (diCenzo *et al*., 2014) zur Aufklärung der Funktionsweise des mehrteiligen Genoms beitragen. Auch ein Austausch der nativen Replikationsursprünge gegen eine heterologe *repABC*-Kassette ließe sich mit Hilfe eines Cre/*lox*-Kassettenaustausches (RMCE) realisieren (Turan *et al*., 2011; Turan *et al*., 2013). Ein pABC-Derivat, das als *repABC*-Kassetten-Donor fungiert, könnte hierbei als *Shuffle*-Vektor zum Austausch eines nativen Replikationsursprungs eingesetzt werden. Die zweischrittige Austausch-Rekombination würde hierbei *in-vivo* und bestenfalls innerhalb eines Zellzyklus stattfinden. Die hohe Effizienz der Rekombinase limitiert diese Anwendung voraussichtlich nicht. Von großem Vorteil sind zudem auf das Cre/*lox*-System adaptierte Resistenzkassetten (P*min2-lox*-R), die eine zuverlässige Selektion auf den vollständigen Austausch ermöglichen. Mit Hilfe der charakterisierten *loxL/R-Spacer*-Mutanten (Kap. 3.1.2) und einer Etablierung des Selektionssystems auf dem Donor-Plasmid, welches sich nach erfolgreichem RMCE beispielsweise via ctRNA-Überexpression deletieren lässt (diCenzo *et al*., 2014), ist auch eine mehrfache Durchführung dieser Strategie möglich. Mit Hilfe des in *S. meliloti* JDSm106 bereitgestellten dualen Replikon-Markierungssystems (Kap. 6.4) könnten in Replikon-Koordinationsstudien schließlich Änderungen der Segregationsdynamik aufgedeckt und wechselseitige Beziehungen zwischen Megaplasmiden und/oder artifiziellen Mini-Replikons hergeleitet werden.

Der modulare Aufbau des Replikationsursprungs (Kap. 2.2.1) erlaubt außerdem die Konstruktion chimärer *repABC*-Transkriptionseinheiten, die eine systematische Untersuchung der intrinsischen Regulation der Replikationsursprünge ermöglichen. Dabei könnte eine „Modul"-spezifische Segregationsdynamik zutage treten, die sich auf P*repABC*, die intergenische Region zwischen *repB* und *repC*, oder die kodierenden Sequenzen *repAB* und *repC* zurückführen ließe. Freilich darf die Rolle unbekannter, chromosomaler Faktoren nicht außer Acht gelassen werden. Die Identifizierung einer Region, die Einfluss auf den Segregationsstart hat und etwa das GAATCGG-Motiv trägt (Kap. 3.2.1.2), könnte wertvolle Hinweise auf eine Verbindung der *repABC*-Regulation zu zentralen Regulatoren des Zellzyklus liefern (Kap. 3.2.1.2). Neben einer regulatorischen DNA-Sequenz, die beispielsweise in *V. cholerae* im Zuge der DNA-Replikation dupliziert wird, dabei den Repressor des sekundärchromosomalen Replikationsinitiators bindet und seine zytosolische Konzentration reduziert (Val *et al*., 2016), könnte auch ParA, das in dem Betaproteobakterium *Burkholderia cenocepacia* als Regulator der extra-chromosomalen Replikons identifiziert wurde (Du *et al*., 2016), eine wichtige Rolle bei der Synchronisierung *repABC*-basierter Replikons mit dem Chromosom spielen.

Artifizielle, genetische Schaltkreise sind ein Beispiel für synthetische Anwendungen, die eine strikte Kopienzahlkontrolle des DNA-Vehikels erfordern (Lee *et al*., 2016). Auch in rhizobiellen Vertretern besteht Bedarf für Expressionsplattformen mit sehr niedriger Kopienzahl (Yip *et al*., 2015), der bislang lediglich in *R. etli* in Form eines IntA-basierten Integrasesystems zur chromosomalen Integration einer Expressionskassette bedacht wurde (Hernández-Tamayo *et al*., 2016). Chromosomale Integrationen sind aufgrund unvorhersehbarer Positionseffekte jedoch grundsätzlich problematisch. Expressionsanalysen zu einer IPTG-induzierten P*lac-eGFP*-Kassette führten in *E. coli* bis zu 300-fach modulierte Expressionslevel herbei, die auf Positions-bedingte, regulatorische Effekte zurückgeführt werden (Bryant *et al*., 2014). Mit pABCa-c (bzw. beliebigen Derivaten mit korrespondierendem oriVSm-Modul) konnte in dieser Arbeit hingegen ein Vektor-Set etabliert werden, das den Aufbau komplexer Expressionssysteme mit definierter MCS in Einzelkopie ermöglicht. Eine stabile Vererbung aller drei Vektoren bei unverändertem Wachstumsverhalten der Zellkultur deutet zudem auf eine vollumfängliche Kompatibilität der artifiziellen Mini-Replikons untereinander sowie zu pSymA und pSymB hin. Die hohe Transformationseffizienz, die selbst einen Co-Transfer aller drei Derivate in *S. meliloti* Rm1021 und somit die simultane Aufnahme von über 25 kb heterologer DNA ermöglicht (Kap. 6.4), ist ein weiteres Alleinstellungsmerkmal und hebt die praktische Anwendbarkeit gegenüber klassischen Expressionsvektoren hevor.

Die pABC Vektorfamilie hat das Potential, technische Limitierungen in *S. meliloti* zu überwinden. pABCs erfüllen erstmals die erforderlichen Voraussetzung, um in Kapitel 3.1.1 diskutierte *Genome Editing*-Methoden in *S. meliloti* zu realisieren. So basieren insbesondere CRISPR/Cas9-Systeme häufig auf einem Mehrkomponentensystem, das ggf. multiple Expressionseinheiten der sgRNA und Nuklease voneinander trennt und ausgeglichene Transkriptionslevel erfordert (Zhou *et al*., 2016) (Kap. 2.1.2). Auch in Hinblick auf synthetische Anwendungen eröffnen pABCs als DNA-Vehikel neue Möglichkeiten. Im BGM-Vektorsystem fungiert das Genom von *B. subtilis* als adäquate Plattform für bis zu 3 Mb große DNA-Moleküle, weshalb sich dieser Organismus für die Assemblierung sehr großer DNA-Konstrukte durchsetzen konnte (Juhas, 2016; Iwata *et al*., 2013; Itaya und Tsuge, 2011; Itaya *et al*., 2008; Yonemura *et al*., 2007). Die Parallelen zu *repABC*-Replikons, die in Form nativer Megaplasmide chromosomalen Umfang erreichen und dabei keinen strengen, architektonischen Regeln zu folgen scheinen (Kap. 3.1.2), sind offenkundig und lassen auf eine hohe Klonierungskapazität der artifiziellen Pendants hoffen. Freilich muss diese Kapazität in weiteren Experimenten evaluiert werden, und GC-Gehalt, DNA-Herkunft bzw. DNA-Modifizierungen, sowie bekannte Symmetrieelemente sollten hierbei Berücksichtigung finden. Dabei könnten systematische Genom-Rearrangement-Studien wichtige Erkenntnisse hinsichtlich

stabilisierender, bzw. Symmetrie- und Struktur-verleihender Elemente liefern (Kap. 3.1.2). Sollten sich pABC Vektoren als geeignete Vehikel sehr großer DNA-Fragmente erweisen, müsste zudem die Stabilität der heterologen DNA evaluiert werden. Die hohe Rekombinationsrate in Chassis-Organismen wie *E. coli* oder *S. cerevisiae* begünstigt einerseits die DNA-Assemblierung (Kap. 2.1.1.1), erschwert andererseits jedoch die stabile Vererbung solcher DNA, die Gen-Duplikationen oder repetitive Sequenzen enthält. Aus diesem Grund wurden unter anderem Rekombinase-defiziente Stämme mit transient induzierbarem Rekombinationssystem entwickelt (Ogawa *et al*., 2015). Die Rekombinationseffizienz in *S. meliloti* hingegen ist vergleichsweise niedrig. Während zur genomischen Integration heterologer Konstrukte mindestens ~220 bp große homologe DNA-Fragmente mit ~80-prozentiger Sequenzähnlichkeit notwendig sind (Becker *et al*., 2009; persönliche Mitteilung Prof. Anke Becker), werden zum Nachweis spontaner Replikon-Fusionen sogar >1000 bp große Homologien vorausgesetzt (Guo *et al*., 2003). In *Bottom-Up*-Ansätzen (Kap. 2.1.2) könnte dies von großem Vorteil sein.

Das größte Potential in Hinblick auf *Genetic Engineering*-Anwendungen liegt jedoch in der Kombination *repABC*-basierter DNA-Vehikel mit einem leistungsfähigen *in-vivo*-Rekombinationssystem. Die ABC *Cloning*-Methode weist gegenüber alternativen *Direct Cloning*-Verfahren entscheidende Vorzüge auf. So bedingt die "Cre/*loxP* plus BAC"-Strategie, mittels derer große Gen-Cluster aus *A. tumefaciens* isoliert wurden (Hu *et al*., 2016), die Aufreinigung und anschließende Elektroporation der isolierten DNA in den Ziel-Organismus. Der DNA-Transfer via Elektroporation oder chemischer Transformation ist jedoch selbst in *E. coli*, das eine vergleichsweise hohe Elektroporationskapazität aufweist, auf wenige hundert kb große DNA-Moleküle limitiert (Sheng *et al*., 1995). Endogene sowie RP2/RP4-basierte Konjugationssysteme hingegen ermöglichen den Transfer ganzer Megaplasmide zwischen rhizobiellen Modellorganismen wie *S. meliloti*, *R. leguminosarum* und *A. tumefaciens* (Kondorosi *et al*., 1982; Quandt *et al*., 2004; Ding und Hynes, 2009; Blanca-Ordóñez *et al*., 2010). Die Möglichkeit des flexiblen Transfers sehr großer DNA-Moleküle stellt insbesondere gegenüber modernen, TAR-basierten *Engineering*-Anwendungen einen fundamentalen Vorteil dar (Kap. 2.1.2). Der konjugale Transfer zwischen *E. coli*, *S. meliloti*, *A. tumefaciens* konnte im Rahmen dieser Arbeit demonstriert werden (Kap. 3.2.1.3). Analog zum ABC *Cloning*-Verfahren bedient sich auch die "*oriT Directed Cloning*"-Strategie einem RK2-basierten Konjugationssystem und bewerkstelligt den effizienten Transfer großer DNA-Fragmente zwischen verschiedenen Gram-negativen Organismen (Kvitko *et al*., 2013; Chain *et al*., 2000). Da hierbei jedoch anstelle eines individuellen Replikons von *oriT*-Sequenzen flankierte DNA-Fragmente mobilisiert werden, ist die Konjugationseffizienz vergleichsweise gering und limitiert die Größe klonierbarer Konstrukte. Diese Einschränkung konnte in einer optimierten Klonierungsstrategie von diCenzo *et. al* (2016) gemindert werden, indem die Region von Interesse vor ihrem Transfer mittels des Flp/*FRT*-

Rekombinasesystems ausgeschnitten wurde. Der zweischrittige Vorgang, der mangels Replikationsfähigkeit des isolierten Fragments innerhalb eines Zellzyklus stattfinden muss, resultierte mit einer Frequenz von ~10^{-7} (pro Donor und Rezipient) in dem Transfer einer 69 kb großen Region nach *E. coli*. Die Cre/*lox*-vermittelte Generierung von pIso-exo hingegen wies eine 10000-fach höhere Frequenz auf (Happel, 2016), und Isolierung und Transfer des autonom replizierenden Mini-Relikons mussten nicht in unmittelbarer Folge vonstatten gehen. Demzufolge ist davon auszugehen, dass die Klonierungskapazität der ABC *Cloning*-Methode deutlich über der alternativer Systeme liegt.

Auch BrET ermöglicht die *in-vivo* Klonierung sehr großer DNA-Konstrukte in *B. subtilis*, die ebenfalls vor ihrem Transfer auf ein replizierendes, lineares Vektor-*Backbone* übertragen werden (Kap. 2.1.2). Abhängig vom verwendeten Replikationsursprung propagiert dieses in *B. subtilis* jedoch mit mindestens 2-3 Kopien pro Zelle (Tanaka und Ogura, 1998). Da HR- und Rekombinase-basierte Folgeprozeduren eine einfache Kopienzahl der Rekombinationssubstrate erfordern, sind anschließende *Engineering*-Anwendungen nur noch eingeschränkt möglich. Im Gegensatz hierzu liegt das Produkt der ABC *Cloning*-Methode in einfacher Kopienzahl vor. Dies ist eine essentielle Voraussetzung, um die isolierte DNA mittels einer weiteren, spezifischen Cre-Rekombination beispielsweise in einen neuen genomischen Lokus zu integrieren. Da im Genom des Donor-Organismus lediglich eine *loxLR*-Sequenz verbleibt, während das Isolationsprodukt eine native *loxP* (und ggf. weitere *lox Spacer*-Mutanten) enthält, sind Cre/*lox*-basierte Folgeprozesse sowohl im Donor, als auch im rhizobiellen Rezipienten durchführbar.

Mit Hilfe einer pABC-basierten Plattform, die großes Potential in der *in-vivo* DNA-Assemblierung verspricht (Kap. 3.2.1.1), einem leistungsfähigen Konjugationssystem zum Transfer großer genomischer Regionen (Kap. 3.2.1.3), sowie verbesserter und erweiterter Cre-Rekombinationsstrategien (Kap. 3.1.2) könnten schließlich ambitionierte *Bottom-Up*-Ansätze zur Etablierung extra-chromosomaler Konstrukte realisiert werden. pABC Vektoren könnten außerdem eine Komplementation essentieller Sequenzen *in-trans* gewährleisten und beispielsweise einen am MEGA-Verfahren (Kap. 2.1.2) orientierten *Top-Down*-Ansatz zur Chromosomen-Reduktion ermöglichen. Neue, innovative *Engineering*-Ansätze profitieren von den individuellen Vorzügen, die *S. meliloti* gegenüber *E. coli*, *B. subtilis* und *S. cerevisiae* aufweist. Dank der im Rahmen dieser Arbeit entwickelten Methoden und Techniken kann das Anwendungspotential eines Alphaproteobakteriums im Bereich der Synthetischen Mikrobiologie somit erstmals zu dem etablierter Chassis-Organismen aufschließen.

4. Fazit

Jahrzehntelange Grundlagenforschung bildet das Fundament der Synthetischen Biologie, die die Enwicklung neuartiger, biologischer Systeme mit einem definierten Anwendungsaspekt zum Ziel hat. *Genetic Engineering* bildet die Basis und bedarf innovativer Strategien zur *in-vivo* Modifikation und Handhabung großer DNA-Moleküle. Möglichst einfach und flexibel anwendbare Techniken, die die Bearbeitung heterologer sowie endogener DNA erlauben, sind hierbei von großem Vorteil (Abb. 4.1).

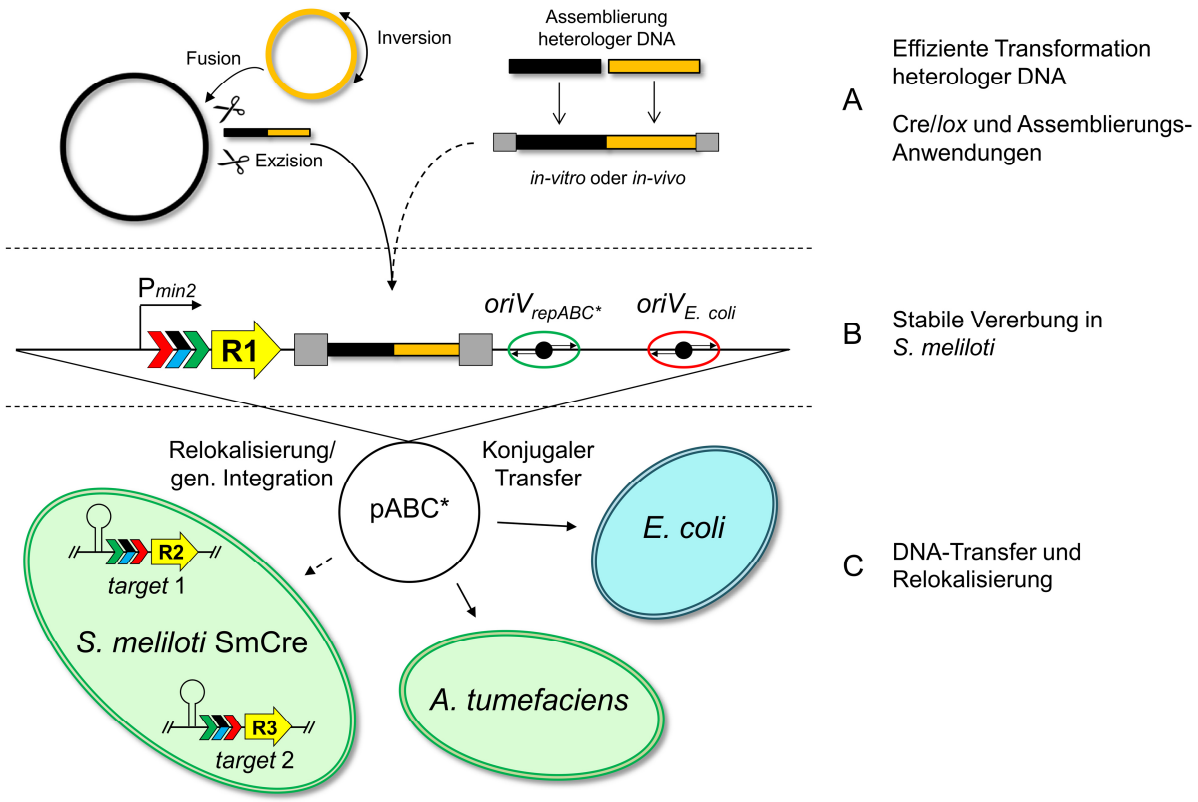


Abbildung 4.1: Generalisierte Darstellung entwickelter *repABC*- und Cre/*lox*-basierter *Engineering*-Strategien und Ausblick auf mögliche Anwendungen (gestrichelte Pfeile). **A:** Die hohe Transformationseffizienz von *S. meliloti* ΔhsdR erleichtert die Einführung großer DNA-Konstrukte. HR- und Cre/*lox*-basierte Rekombinationssysteme bieten zudem die Möglichkeit, artifizielle DNA-Konstrukte zu assemblieren, oder die Genomarchitektur mittels Fusion, Inversion und Deletion in großem Ausmaß zu wandeln. **B:** pABC-Derivate (pABC*), die auf nativen oder induzierbaren *repABC*-Replikationsursprüngen ($oriV_{repABC*}$) basieren, vermitteln in *S. meliloti* die Vererbung heterologer DNA oder der ABC *Cloning*-Produkte in Einzelkopie. *loxL/R*-Sequenzen (dreigliedrige Pfeile, mutiertes Element in rot) mit unterschiedlichen *Spacer*-Motiven (blau/schwarzer Pfeil) ermöglichen in fortentwickelten Cre/*lox*-Systemen spezifische, gerichtete Folgeprozesse. Standardisierte Resistenzkassetten (P*min2*-*lox*-R) gewährleisten eine verlässliche Selektion auf erfolgreiche Rekombinationsereignisse. **C:** pABC*-Derivate könnten mit Hilfe des *S. meliloti*-Stammes SmCre somit die Relokalisierung isolierter oder heterologer DNA in einen neuen genomischen Lokus (*target 1/2*) ermöglichen. Ein RK2-basiertes Mobilisierungssystem vermittelt zudem den horizontalen Transfer dieser Replikons in kompatible Organismen wie *E. coli* und *A. tumefaciens*. Die weite Verbreitung *repABC*-basierter Replikationssysteme (Galardini *et al*., 2013) legt die Kompatibilität *repABC*-basierter Vehikel zu weiteren α-Rhizobien nahe. $oriV_{repABC}$: *repABC*-basierter Replikationsursprung. $oriV_{E.\ coli}$: *E. coli*-kompatibler Replikationsursprung.

S. meliloti stellt als Stickstoff-Fixierer einen biotechnologisch relevanten Organismus mit großem Anwendungspotential im Bereich des G*enetic Engineering* dar. Die technischen Möglichkeiten genetischer Modifikationen waren bislang jedoch schnell erschöpft, sei es aufgrund langwieriger Klonierungsprozeduren oder einem unzureichenden Repertoire genetischer Werkzeuge und Vektoren. Diese Limitierungen konnten in dieser Arbeit mithilfe der CFGE-Methode, optimierter Cre/*lox*-Rekombinationsstrategien, sowie der pABC Vektor-Familie entscheidend verringert werden. Umfangreiche Genom-Umstrukturierungen sowie die simultane Etablierung mehrerer, artifizieller Mini-Replikons bezeugen eine enorme Flexibilität hinsichtlich der Architektur und Organisation des mehrteiligen Genoms und unterstreichen die Eignung von *S. meliloti* als Chassis-Organismus in der Synthetischen Mikrobiologie.

Leistungsfähige *in-vitro* Assemblierungstechniken und sinkende DNA-Synthesekosten begünstigen die *de-novo* Assemblierung sehr großer DNA-Fragmente, die *Bottom-Up*-Ansätze zur (Re-) Konstruktion komplexer Biosynthesewege und artifizieller Chromosome ermöglichen. pABC Vektoren mit Chromosomen-ähnlichen Eigenschaften bergen das Potential, DNA-Moleküle im Umfang mehrerer Mb zu propagieren und als leicht transferierbare DNA-Vehikel für komplexe Expressions- und Assemblierungssysteme zu fungieren. Techniken, die eine flexible Handhabung und Modifikation der DNA *in-vivo* ermöglichen, sind hierbei von großer Bedeutung. ABC *Cloning* kombiniert daher die Idee *repABC*-basierter DNA-Vehikel mit Cre/*lox*-Rekombinationsstrategien und liefert neue methodische Ansätze für biotechnologische Anwendungen (Abb. 4.1 AB). Die Ungewissheit über die tatsächliche Klonierungskapazität und Skalierbarkeit pABC-basierter Klonierungsstrategien verlangt jedoch nach einem besseren Verständnis der Funktionsweise des mehrteiligen Genoms. Neue *Genome Engineering*-Werkzeuge helfen, die zugrundeliegenden Mechanismen aufzuklären und Faktoren zu identifizieren, welche ein *repABC*-basiertes Replikon in den Zellzyklus eines Alphaproteobakteriums integrieren. Von einem tiefgreifenden biologischen Verständnis des mehrteiligen Genoms würden nicht nur Anwendungen in *S. meliloti* profitieren. Die weite Verbreitung des *repABC*-Systems und die robuste Funktion des Cre/*lox*-Systems versprechen auch eine erfolgreiche Adaptation der entwickelten Methoden auf weitere Vertreter der biologisch und biotechnologisch interessanten Klasse der Alphaproteobakterien.

5. Literaturverzeichnis

Abremski, K; Hoess, R (1984). Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. *J Biol Chem. 259* (3), 1509–1514.

Akeno, Y; Ying, BW; Tsuru, S; Yomo, T (2014). A reduced genome decreases the host carrying capacity for foreign DNA. *Microb Cell Fact.* 13 (1), 49.

Albert, H; Dale, EC; Lee, E; Ow, DW (1995). Site-specific integration of DNA into wild-type and mutant *lox* sites placed in the plant genome. *Plant J.* 7 (4), 649–659.

Altenbuchner, J (2016). Editing of the *Bacillus subtilis* Genome by the CRISPR-Cas9 System. *Appl Environ Microbiol.* 82 (17), 5421–5427.

Anderson, JC; Voigt, CA; Arkin, AP (2007). Environmental signal integration by a modular AND gate. *Mol Syst Biol.* 3, 133.

Ara, K; Ozaki, K; Nakamura, K; Yamane, K; Sekiguchi, J; Ogasawara, N (2007). *Bacillus* minimum genome factory: effective utilization of microbial genome information. *Biotechnol Appl Biochem.* 46 (Pt 3), 169–178.

Baba, T; Ara, T; Takai, Y; Okumura, Y; Baba, M; Datsenko, KA *et al*. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol.* 2, 2006.0008.

Bar-Even, A; Noor, E; Lewis, NE.; Milo, R (2010). Design and analysis of synthetic carbon fixation pathways. *Proc Natl Acad Sci U S A.* 107 (19), 8889–8894.

Barrangou, R; van Pijkeren, JP (2016). Exploiting CRISPR-Cas immune systems for genome editing in bacteria 37, 61–68. DOI: 10.1016/j.copbio.2015.10.003.

Bayer, EA; Lamed, R; White, BA; Flint, HJ (2008). From cellulosomes to cellulosomics. *Chem Rec.* 8 (6), 364–377.

Becker, A; Schmidt, M; Jäger, W; Pühler, A (1995). New gentamicin-resistance and lacZ promoter-probe cassettes suitable for insertion mutagenesis and generation of transcriptional fusions. *Gene* 162 (1), 37–39.

Becker, A; Barnett, MJ; Capela, D; Dondrup, M; Kamp, PB; Krol, E; Linke, B; Rüberg, S; Runte, K; Schroeder, B K; Weidner, S; Yurgel, SN; Batut, J; Long, SR; Pühler, A; Goesmann, A. (2009). A portal for rhizobial genomes: RhizoGATE integrates a *Sinorhizobium meliloti* genome annotation update with postgenome data. *J. Biotechnol.*140, 45–50.

Becker, A (2015). Challenges and perspectives in combinatorial assembly of novel exopolysaccharide biosynthesis pathways. *Front Microbiol.* 6, 687.

Bell, JC.; Kowalczykowski, SC (2016). RecA: Regulation and Mechanism of a Molecular Search Engine. *Trends Biochem Sci.* 41 (6), 491–507.

Berg, JM; Stryer, L; Tymoczko, JL (2013). Stryer Biochemie. Springer Spektrum.

Blanca-Ordóñez, H; Oliva-García, JJ; Pérez-Mendoza, D; Soto, MJ; Olivares, J; Sanjuán, J; Nogales, J (2010). pSymA-dependent mobilization of the *Sinorhizobium meliloti* pSymB megaplasmid. *J Bacteriol.* 192 (23), 6309-6312.

Boccard, F; Esnault, E; Valens, M (2005). Spatial arrangement and macrodomain organization of bacterial chromosomes. *Mol Microbiol.* 57 (1), 9–16.

Boch, J; Scholze, H; Schornack, S; Landgraf, A; Hahn, S; Kay, S *et al*., (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326 (5959), 1509–1512.

Bordbar, A; Nagarajan, H; Lewis, NE.; Latif, H; Ebrahim, A; Federowicz, S *et al*., (2014). Minimal metabolic pathway structure is consistent with associated biomolecular interactions. *Mol Syst Biol.* 10, 737.

Bowater, R; Doherty, AJ (2006). Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genet.* 2 (2), e8.

Braff, D; Shis, D; Collins, JJ (2016). Synthetic biology platform technologies for antimicrobial applications. *Adv Drug Deliv Rev.* 105(Pt A), 35-43.

Brennecke, M (2016). Cre-mediated large scale manipulations of the genome in the alphaproteobacterium *Sinorhizobium meliloti*. Masterarbeit, Philipps-Universität Marburg, LOEWE-Zentrum für Synthetische Mikrobiologie, AG Becker.

Broach, JR; Guarascio, VR; Jayaram, M (1982). Recombination within the yeast plasmid 2mu circle is site-specific. *Cell* 29 (1), 227–234.

Brown, WR; Lee, NC; Xu, Z; Smith, MC (2011). Serine recombinases as tools for genome engineering. *Methods.* 53 (4), 372–379.

Bryant, JA; Sellars, LE; Busby, SJ; Lee, DJ (2014). Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Res.* 42 (18), 11383-92.

Cervantes-Rivera, R; Romero-López, C; Berzal-Herranz, A; Cevallos, MA (2010). Analysis of the mechanism of action of the antisense RNA that controls the replication of the *repABC* plasmid p42d. *J Bacteriol*. 192 (13), 3268–3278.

Cervantes-Rivera, R; Pedraza-López, F; Pérez-Segura, G; Cevallos, MA (2011). The replication origin of a *repABC* plasmid. *BMC Microbiol.* 11, 158.

Cevallos, MA.; Cervantes-Rivera, R; Gutiérrez-Ríos, RM (2008). The *repABC* plasmid family. *Plasmid* 60 (1), 19–37.

Chaikind, B; Bessen, JL.; Thompson, DB.; Hu, JH.; Liu, DR. (2016). A programmable Cas9-serine recombinase fusion protein that operates on DNA sequences in mammalian cells. *Nucleic Acids Res.*, DOI: 10.1093/nar/gkw707.

Chain, PS; Hernandez-Lucas, I; Golding, B; Finan, TM (2000). *oriT*-Directed Cloning of Defined Large Regions from Bacterial Genomes: Identification of the *Sinorhizobium meliloti* pExo Megaplasmid Replicator Region. *J Bacteriol.* 182 (19), 5486–5494.

Chan, KM; Liu, YT; Ma, CH; Jayaram, M; Sau, S (2013). The 2 micron plasmid of Saccharomyces cerevisiae: a miniaturized selfish genome with optimized functional competence. *Plasmid* 70 (1), 2–17.

Chandran, V (2013). Type IV secretion machinery: molecular architecture and function. *Biochem Soc Trans.* 41 (1), 17–28.

Charoenpanich, P; Meyer, S; Becker, A; McIntosh, M (2013). Temporal expression program of quorum sensing-based transcription regulation in *Sinorhizobium meliloti*. *J Bacteriol.* 195 (14), 3224–3236.

Chedin, F; Kowalczykowski, SC (2002). A novel family of regulated helicases/nucleases from Gram-positive bacteria: insights into the initiation of DNA recombination. *Mol Microbiol.* 43 (4), 823–834.

Chen, P; Leung, KP (2016). Creation of stable *Pseudomonas aeruginosa* promoter-reporter fusion mutants using linear plasmid DNA transformation. *BMC Res Notes.* 9, 322.

Chen, Q; Fischer, JR; Benoit, VM.; Dufour, NP; Youderian, P; Leong, JM (2008). *In vitro* CpG methylation increases the transformation efficiency of *Borrelia burgdorferi* strains harboring the endogenous linear plasmid lp56. *J Bacteriol.* 190 (24), 7885–7891.

Choi, KR; Lee, SY (2016). CRISPR technologies for bacterial systems: Current achievements and future directions. *Biotechnol Adv.* S0734-9750(16)30103-3.

Citorik, RJ; Mimee, M; Lu, TK (2014). Bacteriophage-based synthetic biology for the study of infectious diseases. *Curr Opin Microbiol.* 19, 59–69.

Cobb, RE; Wang, Y; Zhao, H (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth Biol.* 4 (6), 723–728.

Costantino, N; Court, DL (2003). Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. *Proc Natl Acad Sci U S A.* 100 (26), 15748-53.

Court, DL.; Sawitzke, JA.; Thomason, LC (2002). Genetic engineering using homologous recombination. *Annu Rev Genet.* 36, 361–388.

Dam, P; Kataeva, I; Yang, SJ; Zhou, F; Yin, Y; Chou, W *et al*. (2011). Insights into plant biomass conversion from the genome of the anaerobic thermophilic bacterium *Caldicellulosiruptor bescii* DSM 6725. *Nucleic Acids Res.* 39 (8), 3240–3254.

de Kok, S; Stanton, LH; Slaby, T; Durot, M; Holmes, VF; Patel, KG *et al*., (2014). Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synth Biol.*3 (2), 97–106.

de Nisco, NJ; Abo, RP.; Wu, CM; Penterman, J; Walker, GC (2014). Global analysis of cell cycle gene expression of the legume symbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci U S A.* 111 (9), 3217–3224.

Dekker, J; Marti-Renom, MA.; Mirny, LA. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet.* 14 (6), 390–403.

del Solar, G; Giraldo, R; Ruiz-Echevarria, MJ; Espinosa, M; Diaz-Orejas, R (1998). Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev.*62 (2), 434–464.

Deriano, L; Roth, DB (2013). Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu Rev Genet.* 47, 433–455.

Desjarlais, JR; Berg, JM (1992). Toward rules relating zinc finger protein sequences and DNA binding site preferences. *Proc Natl Acad Sci USA* 89, 7345–7349.

Desjarlais, JR; Berg, JM (1993). Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. *Proc Natl Acad Sci USA* 90, 2256–2260.

diCenzo, G; Milunovic, B; Cheng, J; Finan, TM (2013). The *tRNAarg* gene and *engA* are essential genes on the 1.7-Mb pSymB megaplasmid of *Sinorhizobium meliloti* and were translocated together from the chromosome in an ancestral strain. *J Bacteriol.* 195 (2), 202–212.

diCenzo, GC.; MacLean, AM.; Milunovic, B; Golding, GB; Finan, TM (2014). Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genetic.* 10 (10), e1004742.

diCenzo, G C.; Zamani, M; Milunovic, B; Finan, TM (2016). Genomic resources for identification of the minimal N2 -fixing symbiotic genome. *Environ Microbiol.* 18 (8), 2534–2547.

Ding, H; Hynes, MF (2009). Plasmid transfer systems in the rhizobia. *Can. J. Microbiol.* 55, 917–927.

Döhlemann, J; Brennecke, M; Becker, A (2016). Cloning-free genome engineering in *Sinorhizobium meliloti* advances applications of Cre/*loxP* site-specific recombination. *J Biotechnol.* 10 (233), 160-70.

Donahue, JP; Israel, DA; Peek, RM; Blaser, M J; Miller, GG (2000). Overcoming the restriction barrier to plasmid transformation of *Helicobacter pylori*. *Mol Microbiol.* 37 (5), 1066–1074.

Dong, H; Li, S; Fang, H; Xia, M; Zheng, P; Zhang, D; Sun, J (2016). A newly isolated and identified vitamin B12 producing strain: *Sinorhizobium meliloti* 320. *Bioprocess Biosyst Eng* 39 (10), 1527-37.

Dorman, CJ (2013). Genome architecture and global gene regulation in bacteria: making progress towards a unified model? *Nat Rev Microbiol.* 11 (5), 349–355.

Du, W-L; Dubarry, N; Passot, FM; Kamgoué, A; Murray, H; Lane, D; Pasta F (2016). Orderly Replication and Segregation of the Four Replicons of Burkholderia cenocepacia J2315. *PLoS Genet.* 12 (7), e1006172.

Dubnau, D (1991). Genetic competence in *Bacillus subtilis*. *Microbiol Rev.* 55 (3), 395–424.

Durante-Rodríguez, G; de Lorenzo, V; Martínez-García, E (2014). The Standard European Vector Architecture (SEVA) plasmid toolkit. *Methods Mol Biol.* 1149, 469-78.

Echols, H (1970). Integrative and excisive recombination by bacteriophage lambda: evidence for an excision-specific recombination protein. *J Mol Biol.* 47 (3), 575–583.

Elowitz, MB; Leibler, S (2000). A synthetic oscillatory network of transcriptional regulators. *Nature* 403 (6767), 335–338.

Esnault, E; Valens, M; Espeli, O; Boccard, F (2007). Chromosome structuring limits genome plasticity in *Escherichia coli*. *PLoS Genet.* 3 (12), e226.

Esseling, JJ; Lhuissier, FG; Emons, AM (2003). Nod factor-induced root hair curling: continuous polar growth towards the point of nod factor application. *Plant Physiol.* 132 (4), 1982–1988.

Ettema, TJ; Andersson, SG (2009). The alpha-proteobacteria: the Darwin finches of the bacterial world. *Biol Lett.* 5 (3), 429–432.

Ferri, L; Gori, A; Biondi, EG.; Mengoni, A; Bazzicalupo, M (2010). Plasmid electroporation of *Sinorhizobium strains*: The role of the restriction gene *hsdR* in type strain Rm1021. *Plasmid* 63 (3), 128–135.

Ferullo, DJ; Cooper, DL; Moore, HR; Lovett, ST (2009). Cell cycle synchronization of *Escherichia coli* using the stringent response, with fluorescence labeling assays for DNA content and replication. *Methods* 48 (1), 8–13.

Feschenko, VV; Rajman, LA; Lovett, ST (2003). Stabilization of perfect and imperfect tandem repeats by single-strand DNA exonucleases. *Proc. Natl. Acad. Sci. USA* 100, 1134-1139.

Fleischmann, RD; Adams, MD; White, O; Clayton, RA; Kirkness, EF; Kerlavage, AR *et al*. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd 269. *Science* 269 (5223), 496–512.

Forster, AC; Curch, GM (2006). Towards synthesis of a minimal cell. *Mol Syst Biol.* 2, 45.

Frage, B; Döhlemann, J; Robledo, M; Lucena, D; Sobetzko, P; Graumann, PL.; Becker, A (2016). Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. *Mol Microbiol.* 100 (5), 808-23.

Friedland, AE; Lu, TK; Wang, X; Shi, D; Church, G; Collins, JJ (2009). Synthetic gene networks that count. *Science* 324 (5931), 1199–1202.

Frith, MC; Saunders, NF; Kobe, B; Bailey, TL (2008). Discovering sequence motifs with arbitrary insertions and deletions. *PLoS Comput Biol.* 4 (4), e1000071.

Fung, E; Wong, WW; Suen, JK; Bulter, T; Lee, S; Liao, JC (2005). A synthetic gene-metabolic oscillator. *Nature* 435 (7038), 118–122.

Gaj, T; Sirk, SJ; Barbas, CF 3rd (2014). Expanding the scope of site-specific recombinases for genetic and metabolic engineering. *Biotechnol Bioeng.* 111 (1), 1–15.

Galardini, M; Pini, F; Bazzicalupo, M; Biondi, EG; Mengoni, A (2013). Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. *Genome Biol Evol.* 5 (3), 542–558.

Galibert, F; Finan, TM; Long, SH; Pühler, A; Abola, P; Ampe, F; Barloy-Hubler, F *et al*., (2001). The Composite Genome of the Legume Symbiont *Sinorhizobium meliloti*. *Science* 293 (5530), 663–668.

Gardner, TS; Cantor, CR; Collins, JJ (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403 (6767), 339–342.

Gerdes, K; Howard, M; Szardenings, F (2010). Pushing and pulling in prokaryotic DNA segregation. *Cell* 141 (6), 927–942.

Geurts, R; Fedorova, E; Bisseling, T (2005). Nod factor signaling genes and their function in the early stages of *Rhizobium* infection. *Curr Opin Plant Biol.* 8 (4), 346–352.

Gibson, DG; Benders, GA; Andrews-Pfannkoch, C; Denisova, EA; Baden-Tillson, H; Zaveri, J *et al*., (2008). Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 319 (5867), 1215–1220.

Gibson, KE; Kobayashi, H; Walker, GC (2008). Molecular determinants of a symbiotic chronic infection. *Annu Rev Genet.* 42, 413–441.

Gibson, DG; Young, L; Chuang, RY; Venter, JC; Hutchison, CA 3rd; Smith, HO (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 6 (5), 343–345.

Gibson, DG; Glass, JI; Lartigue, C; Noskov, VN; Chuang, RY; Algire, MA; Benders GA (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* 329 (5987), 52–56.

Gil, R; Silva, FJ; Peretó, J; Moya, A (2004). Determination of the core of a minimal bacterial gene set. *Microbiol Mol Biol Rev.*(3), 518-37.

González, V; Santamaría, RI; Bustos, P; Hernández-González, I; Medrano-Soto, A; Moreno-Hagelsieb, G; Janga, SC *et al*., (2006). The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci U S A* 103 (10), 3834–3839.

Grindley, ND; Whiteson, KL; Rice, PA (2006). Mechanisms of site-specific recombination. *Annu Rev Biochem.*75, 567–605.

Guo, X; Flores, M; Mavingui, P; Fuentes, SI; Hernandez, G; Dávila, G; Palacios, R (2003). Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. *Genome Res.* 13 (8), 1810–1817.

Gust, B; Chandra, G; Jakimowicz, D; Yuqing, T; Bruton, CJ; Chater, Keith F (2004). Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. *Adv Appl Microbiol.* 54, 107–128.

Guzman, LM; Belin, D; Carson, MJ; Beckwith, J (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol*. 177 (14), 4121–4130.

Happel, C (2016). *In vivo* DNA assembly in *Sinorhizobium meliloti* and interspecies DNA transfer. Masterarbeit, Philipps-Universität Marburg, LOEWE-Zentrum für Synthetische Mikrobiologie, AG Becker.

Harrison, CL; Crook, MB; Peco, G; Long, SR; Griffitts, JS (2011). Employing site-specific recombination for conditional genetic analysis in *Sinorhizobium meliloti*. *Appl Environ Microbiol.* 77 (12), 3916–3922.

Harrison, PW; Lower, RP; Kim, NK; Young, JP (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol.* 18 (4), 141–148.

Haykinson, MJ; Johnson, LM; Soong, J; Johnson, RC (1996). The Hin dimer interface is critical for Fis-mediated activation of the catalytic steps of site-specific DNA inversion. *Curr Biol.* 6 (2), 163–177.

He, F; Murabito, E; Westerhoff, HV (2016). Synthetic biology and regulatory networks: where metabolic systems biology meets control engineering. *J R Soc Interface.* 13 (117).

Hendrickson, H; Lawrence, JG (2006). Selection for chromosome architecture in bacteria. *J Mol Evol.* 62 (5), 615–629.

Hernández-Tamayo, R; Torres-Tejerizo, G; Brom, S; Romero, D (2016). Site-specific bacterial chromosome engineering mediated by IntA integrase from *Rhizobium etli*. *BMC Microbiol.*16 (1), 133.

Herridge, DF; Peoples, MB; Boddey, RM (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant Soil.* 311 (1), 1–18.

Hossain, MA.; Barrow, JJ; Shen, Y; Haq, MI; Bungert, J (2015). Artificial zinc finger DNA binding domains: versatile tools for genome engineering and modulation of gene expression. *J Cell Biochem.* 116 (11), 2435–2444.

Hu, S; Liu, Z; Zhang, X; Zhang, G; Xie, Y; Ding, X *et al*., (2016). "Cre/loxP plus BAC": a strategy for direct cloning of large DNA fragment and its applications in *Photorhabdus luminescens* and *Agrobacterium tumefaciens*. *Sci Rep.* 6, 29087.

Hübner, P; Willison, JC; Vignais, PM; Bickle, TA (1991). Expression of regulatory nif genes in *Rhodobacter capsulatus*. *J Bacteriol.* 173 (9), 2993–2999.

Hutchison, CA 3rd; Chuang, RY; Noskov, VN; Assad-Garcia, N; Deerinck, TJ; Ellisman, MH;Gill, J *et al*., (2016). Design and synthesis of a minimal bacterial genome. *Science* 351 (6280), aad6253.

Hynes, MF; Quandt, J; O'Connell, MP; Pühler, A (1989). Direct selection for curing and deletion of *Rhizobium* plasmids using transposons carrying the *Bacillus subtilis sacB* gene. *Gene* 78 (1), 111–120.

Wang, HH; Isaacs, FJ; Carr, PA; Sun, ZZ; Xu, G; Forest, CR; Church, GM (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature.* 460 (7257), 894–898.

Ishino, Y; Shinagawa, H; Makino, K; Amemura, M; Nakata, A (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol.* 169 (12), 5429–5433.

Itaya, M (1999). Effective Cloning of Unmarked DNA Fragments in the *Bacillus subtilis* 168 Genome. *Biosci Biotechnol Biochem.* 63 (3), 602–604.

Itaya, M; Tanaka, T (1990). Gene-directed mutagenesis on the chromosome of *Bacillus subtilis* 168. *Mol Gen Genet.* 223 (2), 268–272.

Itaya, M; Tanaka, T (1997a). Experimental surgery to create subgenomes of *Bacillus subtilis* 168. *Proc Natl Acad Sci U S A* 94 (10), 5378–5382.

Itaya, M; Tanaka, T (1997b). Predicted and Unsuspected Alterations of the Genomes Structure of Genetically Defined *Bacillus subtilis* 168 Strains. *Biosci., Biotechnol., Biochem.* 61 (1), 56–64.

Itaya, M; Fujita, K; Kuroki, A; Tsuge, K (2008). Bottom-up genome assembly using the *Bacillus subtilis* genome vector. *Nat methods* 5 (1), 41–43.

Itaya, M; Tsuge, K (2011). Construction and manipulation of giant DNA by a genome vector. *Methods Enzymol.* 498, 427–447.

Itaya, M; Tsuge, K; Koizumi, M; Fujita, K (2005). Combining two genomes in one cell: stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc Natl Acad Sci U S A* 102 (44), 15971–15976.

Iwata, T; Kaneko, S; Shiwa, Y; Enomoto, T; Yoshikawa, H; Hirota, J (2013). *Bacillus subtilis* genome vector-based complete manipulation and reconstruction of genomic DNA for mouse transgenesis. *BMC Genomics* 14, 300.

Jayaram, M; Ma, CH; Kachroo, AH; Rowley, PA; Guga, P; Fan, HF; Voziyanov, Y (2015). An Overview of Tyrosine Site-specific Recombination: From an Flp Perspective. *Microbiol Spectr.* 3 (4).

Jiang, W; Bikard, D; Cox, D; Zhang, F; Marraffini, LA (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol.*31 (3), 233–239.

Jiang, Y Chen, B; Duan, C; Sun, B; Yang, J; Yang, S (2015). Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol.* 81 (7), 2506–2514.

Jinek, M; Chylinski, K; Fonfara, I; Hauer, M; Doudna, J A.; Charpentier, E (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337 (6096), 816–821.

Jones, KM; Kobayashi, H; Davies, BW; Taga, ME; Walker, GC (2007). How rhizobial symbionts invade plants: the *Sinorhizobium*-Medicago model. *Nat Rev Microbiol.* 5 (8), 619–633.

Juhas, M (2016). On the road to synthetic life: the minimal cell and genome-scale engineering. *Crit Rev Biotechnol.* 36 (3), 416–423.

Kahng, LS; Shapiro, L (2001). The CcrM DNA methyltransferase of *Agrobacterium tumefaciens* is essential, and its activity is cell cycle regulated. *J Bacteriol.* 183 (10), 3065–3075.

Kaneko, S; Akioka, M; Tsuge, K; Itaya, M (2005). DNA shuttling between plasmid vectors and a genome vector: systematic conversion and preservation of DNA libraries using the *Bacillus subtilis* genome (BGM) vector. *J Mol Biol.* 349 (5), 1036–1044.

Karas, BJ; Jablanovic, J; Sun, L; Ma, L; Goldgof, GM; Stam, J *et al*., (2013). Direct transfer of whole genomes from bacteria to yeast. *Nat methods* 10 (5), 410–412.

Karlinsey, JE (2007). lambda-Red genetic engineering in *Salmonella enterica* serovar Typhimurium. *Methods Enzymol.* 421, 199–209.

Ke, R; Mignardi, M; Hauling, T; Nilsson, M (2016). Fourth Generation of Next-Generation Sequencing Technologies: Promise and Consequences. *Hum Mutat*., DOI: 10.1002/humu.23051.

Khan, SR; Gaines, J; Roop, RM; Farrand, SK (2008). Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl Environ Microbiol.* 74 (16), 5053–5062.

Kis, Z; Pereira, HS; Homma, T; Pedrigi, RM; Krams, R (2015). Mammalian synthetic biology: emerging medical applications. *J R Soc Interface.* 12 (106).

Kobayashi, H; Simmons, LA; Yuan, DS; Broughton, WJ; Walker, GC (2008). Multiple Ku orthologues mediate DNA non-homologous end-joining in the free-living form and during chronic infection of *Sinorhizobium meliloti*. *Mol Microbiol.* 67 (2), 350–363.

Kondorosi, A; Kondorosi, E; Pankhurst, CE; Broughton, WJ; Banfalvi, Z (1982). Mobilization of a *Rhizobium meliloti* megaplasmid carrying nodulation and nitrogen fixation genes into other rhizobia and *Agrobacterium*. *Mol Gen Genet.* 188 (3), 433-439.

Koper, P; Żebracki, K; Marczak, M; Skorupska, A; Mazur, A (2016). RepB proteins of the multipartite *Rhizobium leguminosarum* bv. *trifolii* genome discriminate between centromere-like *parS* sequences for plasmid segregational stability. *Mol Microbiol*., DOI: 10.1111/mmi.13472.

Kosuri, S; Church, GM (2014). Large-scale de novo DNA synthesis: technologies and applications. *Nat Methods.* 11 (5), 499–507.

Kouprina, N; Larionov, V (2016). Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma* 125 (4), 621–632.

Kuo, CH; Moran, NA; Ochman, H (2009). The consequences of genetic drift for bacterial genome complexity. *Genome Res*. 19 (8), 1450–1454.

Kurokawa, M; Seno, S; Matsuda, H; Ying, BW (2016). Correlation between genome reduction and bacterial growth. *DNA Res.*, DOI: 10.1093/dnares/dsw035.

Kvitko, BH; McMillan, IA; Schweizer, HP (2013). An improved method for *oriT*-directed cloning and functionalization of large bacterial genomic regions. *Appl Environ Microbiol.* 79 (16), 4869–4878.

Kwak, J; Jiang, H; Kendrick, KE (2002). Transformation using in vivo and in vitro methylation in *Streptomyces griseus*. *FEMS Microbiol Lett.* 209 (2), 243–248.

Langer, SJ; Ghafoori, AP; Byrd, M; Leinwand, L (2002). A genetic screen identifies novel non-compatible *loxP* sites. *Nucleic Acids Res.* 30 (14), 3067–3077.

Larionov, V; Kouprina, N; Graves, J; Resnick, MA (1996). Highly selective isolation of human DNAs from rodent-human hybrid cells as circular yeast artificial chromosomes by transformation-associated recombination cloning. *Proc Natl Acad Sci U S A* 93 (24), 13925–13930.

Lartigue, C; Glass, JI; Alperovich, N; Pieper, R; Parmar, PP; Hutchison, CA 3rd *et al*. (2007). Genome transplantation in bacteria: changing one species to another. *Science* 317 (5838), 632–638.

Lartigue, C; Vashee, S; Algire, MA; Chuang, RY; Benders, GA; Ma, L; Noskov, VN *et al*. (2009). Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* 325 (5948), 1693–1696.

Lau, IF; Filipe, SR; Søballe, B; Økstad, OA; Barre, FX; Sherratt, DJ (2003). Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol.* 49 (3), 731–743.

Lee, JW; Gyorgy, A; Cameron, DE; Pyenson, N; Choi, KR; Way, JC; Silver PA *et al*. (2016). Creating Single-Copy Genetic Circuits. *Mol Cell.* 63 (2), 329-36.

Lee, DJ; Bingle, LE; Heurlier, K; Pallen, MJ; Penn, CW; Busby, SJ; Hobman, JL (2009). Gene doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains. *BMC Microbiol.* 9, 252.

Lee, G; Saito, I (1998). Role of nucleotide sequences of *loxP* spacer region in Cre-mediated recombination. *Gene* 216 (1), 55–65.

Levy, A; Goren, MG; Yosef, I; Auster, O; Manor, M; Amitai, G *et al*., (2015). CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* 520 (7548), 505–510.

Li, Y; Zhu, X; Zhang, X; Fu, J; Wang, Z; Chen, T; Zhao, X (2016). Characterization of genome-reduced *Bacillus subtilis* strains and their application for the production of guanosine and thymidine. *Microb Cell Fact.* 15, 94.

Liu, Z; Xie, Y; Zhang, X; Hu, X; Li, Y; Ding, X *et al*., (2016). Efficient Construction of Large Genomic Deletion in *Agrobacterium tumefaciens* by Combination of Cre/*loxP* System and Triple Recombineering. *Curr Microbiol.* 72 (4), 465–472.

Livak, KJ; Schmittgen, TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 25 (4), 402-8.

Lobry, JR; Louarn, JM (2003). Polarisation of prokaryotic chromosomes. *Curr Opin Microbiol.* 6 (2), 101–108.

Lynd, LR; Laser, MS; Bransby, D; Dale, BE.; Davison, B; Hamilton, R *et al*., (2008). How biotech can transform biofuels. *Nat Biotechnol.* 26 (2), 169–172..

MacDonald, IC; Deans, TL (2016). Tools and applications in synthetic biology. *Adv Drug Deliv Rev.* 105(Pt A),20-34.

MacLellan, SR; Zaheer, R; Sartor, AL; MacLean, AM; Finan, TM. (2006). Identification of a megaplasmid centromere reveals genetic structural diversity within the *repABC* family of basic replicons. *Mol Microbiol.* 59 (5), 1559–1575.

Makarova, KS; Wolf, YI; Koonin, EV (2009). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol Direct.* 4, 19.

Malgieri, G; Palmieri, M; Russo, L; Fattorusso, R; Pedone, PV; Isernia, C (2015). The prokaryotic zinc-finger: structure, function and comparison with the eukaryotic counterpart. *FEBS J.* 282 (23), 4480–4496.

Manabe, K; Kageyama, Y; Morimoto, T; Ozawa, T; Sawada, K; Endo, K *et al*., (2011). Combined effect of improved cell yield and increased specific productivity enhances recombinant enzyme production in genome-reduced *Bacillus subtilis* strain MGB874. *Appl Environ Microbiol.* 77 (23), 8370–8381.

Marinus, MG (2010). DNA Methylation and Mutator Genes in Escherichia coli K-12. *Mutat Res.* 705 (2), 71–76.

Mavingui, P; Flores, M; Guo, X; Dávila, G; Perret, X; Broughton, WJ; Palacios, R (2002). Dynamics of genome architecture in *Rhizobium* sp. strain NGR234. *J Bacteriol.*184 (1), 171–176.

Mazur, A; Koper, P (2012). Rhizobial plasmids — replication, structure and biological role. *Cent Eur J Biol.* 7 (4).

Mazur, A; Majewska, B; Stasiak, G; Wielbo, J; Skorupska, A (2011). *repABC*-based replication systems of *Rhizobium leguminosarum* bv. *trifolii* TA1 plasmids: incompatibility and evolutionary analyses. *Plasmid* 66 (2), 53–66.

McAnulla, C; Edwards, A; Sanchez-Contreras, M; Sawers, RG; Downie, JA (2007). Quorum-sensing-regulated transcriptional initiation of plasmid transfer and replication genes in *Rhizobium leguminosarum* biovar *viciae*. *Microbiology* 153 (Pt 7), 2074–2082.

McCutcheon, JP; Moran, NA (2012). Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol.*10 (1), 13–26.

Mergaert, P; Uchiumi, T; Alunni, B; Evanno, G; Cheron, A; Catrice, O *et al*., (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. *Proc Natl Acad Sci U S A* 103 (13), 5230–5235.

Milunovic, B; diCenzo, GC; Morton, RA; Finan, TM (2014). Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of *Sinorhizobium meliloti*. *J Bacteriol.* 196 (4), 811–824.

Missirlis, PI; Smailus, DE; Holt, RA (2006). A high-throughput screen identifying sequence and promiscuity characteristics of the *loxP* spacer region in Cre-mediated recombination. *BMC Genomics* 7, 73.

Moënne-Loccoz, Y; Baldani, JI; Weaver, RW (1995). Sequential heat-curing of Tn5-Mob-*sac* labelled plasmids from *Rhizobium* to obtain derivatives with various combinations of plasmids and no plasmid. *Lett Appl Microbiol.* 20 (3), 175–179.

Mojica, FJ; Montoliu, L (2016). On the Origin of CRISPR-Cas Technology: From Prokaryotes to Mammals. *Trends Microbiol.* 24 (10), 811-20.

Moreno, E (1998). Genome evolution within the alpha Proteobacteria: why do some bacteria not possess plasmids and others exhibit more than one different chromosome? *FEMS Microbiol Rev.* 22 (4), 255–275.

Moscou, MJ; Bogdanove, AJ (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* 326 (5959), 1501.

Murray, NE (2000). Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol Mol Biol Rev.* 64 (2), 412–434.

Mus, F; Crook, MB; Garcia, K; Garcia Costas, A; Geddes, BA; Kouri, ED; Paramasivan, P *et al*., (2016). Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes. *Appl Environ Microbiol.* 82 (13), 3698–3710.

Nafissi, N; Slavcev, R (2014). Bacteriophage recombination systems and biotechnical applications. *Appl Microbiol Biotechnol.* 98 (7), 2841–2851.

Nemhauser, JL; Torii, KU (2016). Plant synthetic biology for molecular engineering of signalling and development. *Nat Plants.* 2, 16010.

Ng, CY; Khodayari, A; Chowdhury, A; Maranas, CD (2015). Advances in de novo strain design using integrated systems and synthetic biology tools. *Curr Opin Chem Biol.* 28, 105–114.

Nikel, PI; Chavarria, M; Danchin, A; de Lorenzo, V (2016). From dirt to industrial applications: *Pseudomonas putida* as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr Opin Chem Biol.* 34, 20–29.

Oehler, S; Amouyal, M; Kolkhof, P; von Wilcken-Bergmann, B; Müller-Hill, B (1994). Quality and position of the three lac operators of *E. coli* define efficiency of repression. *EMBO J.* 13 (14), 3348–3355.

Ogawa, T; Iwata, T; Kaneko, S; Itaya, M; Hirota, J (2015). An inducible *recA* expression *Bacillus subtilis* genome vector for stable manipulation of large DNA fragments. *BMC Genomics* 16, 209.

Oh, JH; van Pijkeren, JP (2014). CRISPR-Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucleic Acids Res.* 42 (17), e131. DOI: 10.1093/nar/gku623.

Oldenburg, KR; Vo, KT; Michaelis, S; Paddon, C (1997). Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* 25 (2), 451–452.

Oresnik, IJ; Liu, SL; Yost, CK; Hynes, MF (2000). Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. *J Bacteriol.* 182 (12), 3582–3586.

Orosz, A; Boros, I; Venetianer, P (1991). Analysis of the complex transcription termination region of the *Escherichia coli rrnB* gene. *Eur J Biochem.* 201 (3), 653–659.

Ow, DW (2016). The long road to recombinase-mediated plant transformation. *Plant Biotechnol J.* 14 (2), 441–447.

Padilla-Vaca, F; Anaya-Velázquez, F; Franco, B (2015). Synthetic biology: Novel approaches for microbiology. *Int Microbiol.* 18 (2), 71–84.

Pandey, RP; Parajuli, P; Koffas, M A; Sohng, JK (2016). Microbial production of natural and non-natural flavonoids: Pathway engineering, directed evolution and systems/synthetic biology. *Biotechnol Adv.* 34 (5), 634–662.

Pappas, KM; Winans, SC (2003a). A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. *Mol Microbiol.* 48 (4), 1059–1073.

Pappas, KM; Winans, SC (2003b). The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. *Mol Microbiol*. 49 (2), 441–455.

Pérez-Oseguera, A; Cevallos, MA (2013). RepA and RepB exert plasmid incompatibility repressing the transcription of the *repABC* operon. *Plasmid* 70 (3), 362–376.

Pini, F, Galardini, M, Bazzicalupo, M, Mengoni, A (2011). Plant-bacteria association and symbiosis: are there common genomic traits in alphaproteobacteria? *Genes* 2 (4), 1017–1032.

Pinto, UM; Flores-Mireles, AL; Costa, ED; Winans, SC (2011). RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within *repC* and functions only *in cis*. *Mol Microbiol.* 81 (6), 1593–1606.

Pinto, UM; Pappas, KM; Winans, SC (2012). The ABCs of plasmid replication and segregation. *Nat Rev Microbiol.* 10 (11), 755–765.

Pósfai, G; Plunkett, G 3rd; Fehér, T; Frisch, D; Keil, GM; Umenhoffer, K *et al*., (2006). Emergent properties of reduced-genome *Escherichia coli*. *Science* 312 (5776), 1044–1046.

Poteete, AR (2001). What makes the bacteriophage lambda Red system useful for genetic engineering: molecular mechanism and biological function. *FEMS Microbiol Lett.* 201 (1), 9–14.

Pyne, ME; Moo-Young, M; Chung, DA; Chou, CP (2015). Coupling the CRISPR/Cas9 System with Lambda Red Recombineering Enables Simplified Chromosomal Gene Replacement in *Escherichia coli*. *Appl Environ Microbiol.* 81 (15), 5103–5114.

Quandt, J; Clark, RG; Venter, AP; Clark, SR; Twelker, S; Hynes MF (2004). Modified RP4 and Tn5-Mob derivatives for facilitated manipulation of large plasmids in Gram-negative bacteria. *Plasmid* 52 (1), 1-12.

Rajewska, M; Wegrzyn, K; Konieczny, I (2012). AT-rich region and repeated sequences - the essential elements of replication origins of bacterial replicons. *FEMS Microbiol Rev*. 36 (2), 408–434.

Ramírez-Romero, MA; Soberón, N; Pérez-Oseguera, A; Téllez-Sosa, J; Cevallos, MA (2000). Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. *J Bacteriol*. 182 (11), 3117–3124.

Ramírez-Romero, MA; Téllez-Sosa, J; Barrios, H; Pérez-Oseguera, A; Rosas, V; Cevallos, MA (2001). RepA negatively autoregulates the transcription of the *repABC* operon of the *Rhizobium etli* symbiotic plasmid basic replicon. *Mol Microbiol*. 42 (1), 195–204.

Rivera-Urbalejo, A; Pérez-Oseguera, Á; Carreón-Rodríguez, OE; Cevallos, MA (2015). Mutations in an antisense RNA, involved in the replication control of a *repABC* plasmid, that disrupt plasmid incompatibility and mediate plasmid speciation. *Plasmid* 78, 48–58.

Rockström, J; Steffen, W; Noone, K; Persson, A; Chapin, FS 3rd; Lambin, EF *et al*., (2009). A safe operating space for humanity. *Nature* 461 (7263), 472–475.

Ruan, Y; Zhu, L; Li, Q (2015). Improving the electro-transformation efficiency of *Corynebacterium glutamicum* by weakening its cell wall and increasing the cytoplasmic membrane fluidity. *Biotechnol Lett.* 37 (12), 2445–2452.

Sadykov, MR (2016). Restriction-Modification Systems as a Barrier for Genetic Manipulation of *Staphylococcus aureus*. *Methods Mol Biol.* 1373, 9–23.

Sallström, B; Andersson, SG (2005). Genome reduction in the alpha-Proteobacteria. *Curr Opin Microbiol*. 8 (5), 579–585.

Schäfer, A; Tauch, A; Jäger, W; Kalinowski, J; Thierbach, G; Pühler, A (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 145 (1), 69–73.

Schlüter, JP; Reinkensmeier, J; Barnett, MJ; Lang, C; Krol, E; Giegerich, R; Becker, A (2013). Global mapping of transcription start sites and promoter motifs in the symbiotic alpha-proteobacterium *Sinorhizobium meliloti* 1021. *BMC Genomics* 14, 156.

Schumacher, MA (2012). Bacterial plasmid partition machinery: a minimalist approach to survival. *Curr Opin Struct Biol.* 22 (1), 72–79.

Schwander, T; Schada von Borzyskowski, L; Burgener, S; Cortina, NS; Erb, TJ (2016). A synthetic pathway for the fixation of carbon dioxide *in vitro*. *Science* 354 (6314), 900-904.

Schwartz, MA; Shapiro, L (2011). An SMC ATPase mutant disrupts chromosome segregation in *Caulobacter*. *Mol Microbiol.* 82 (6), 1359–1374.

Senecoff, JF; Rossmeissl, PJ; Cox, MM (1988). DNA recognition by the FLP recombinase of the yeast 2 mu plasmid. A mutational analysis of the FLP binding site. *J Mol Biol.* 201 (2), 405–421.

Shalem, O; Sanjana, NE;Zhang F (2015). High-throughput functional genomics using CRISPR–Cas9. *Nat Rev Genet.* 16 (5): 299-311.

Sheng, Y; Mancino, V; Birren, B (1995). Transformation of *Escherichia coli* with large DNA molecules by electroporation. *Nucleic Acids Res.* 23 (11), 1990-1996.

Sherratt, DJ; Søballe, B; Barre, FX; Filipe, S; Lau, I; Massey, T; Yates, J (2004). Recombination and chromosome segregation. *Philos Trans R Soc Lond B Biol Sci.* 359 (1441), 61–69.

Shuman, S; Glickman, MS (2007). Bacterial DNA repair by non-homologous end joining. *Nat Rev Microbiol.*5 (11), 852–861.

Sibley, CD; MacLellan, SR; Finan, T (2006). The *Sinorhizobium meliloti* chromosomal origin of replication. *Microbiology* 152 (Pt 2), 443–455.

Slager, J; Veening, JW (2016). Hard-Wired Control of Bacterial Processes by Chromosomal Gene Location. *Trends Microbiol*. 24 (10), 788-800.

Soberón, N; Venkova-Canova, T; Ramírez-Romero, MA; Téllez-Sosa, J; Cevallos, MA (2004). Incompatibility and the partitioning site of the *repABC* basic replicon of the symbiotic plasmid from *Rhizobium etli*. *Plasmid* 51 (3), 203–216.

Sonoda, E; Hochegger, H; Saberi, A; Taniguchi, Y; Takeda, S (2006). Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair* 5 (9-10), 1021–1029.

Sternberg, N; Hamilton, D; Austin, S; Yarmolinsky, M; Hoess, R (1981). Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harb Symp Quant Biol.* 45 Pt 1, 297–309.

Suzuki, H; Takahashi, S; Osada, H; Yoshida, K (2011). Improvement of transformation efficiency by strategic circumvention of restriction barriers in *Streptomyces griseus*. *J Microbiol Biotechnol.* 21 (7), 675–678.

Tabata, S; Hooykaas, PJ; Oka, A (1989). Sequence determination and characterization of the replicator region in the tumor-inducing plasmid pTiB6S3. *J Bacteriol.* 171 (3), 1665–1672.

Tallaksen, J; Bauer, F; Hulteberg, C; Reese, M; Ahlgren, S (2015). Nitrogen fertilizers manufactured using wind power: greenhouse gas and energy balance of community-scale ammonia production. *J Cleaner Prod.* 107, 626–635.

Tanaka, T; Ogura, M (1998). A novel *Bacillus natto* plasmid pLS32 capable of replication in *Bacillus subtilis*. *FEBS Lett.* 422 (2), 243–246.

Touzain, F; Petit, MA; Schbath, S; El Karoui, M (2011). DNA motifs that sculpt the bacterial chromosome. *Nat Rev Microbiol.* 9 (1), 15–26.

Toya, Y; Hirasawa, T; Morimoto, T; Masuda, K; Kageyama, Y; Ozaki, K *et al*., (2014). 13 C-metabolic flux analysis in heterologous cellulase production by *Bacillus subtilis* genome-reduced strain. *J Biotechnol.*179, 42–49.

Turan, S; Bode, J (2011). Site-specific recombinases: from tag-and-target- to tag-and-exchange-based genomic modifications. *FASEB J.* 25 (12), 4088–4107.

Turan, S; Galla, M; Ernst, E; Qiao, J; Voelkel, C; Schiedlmeier, B *et al*., (2011). Recombinase-mediated cassette exchange (RMCE): traditional concepts and current challenges. *J Mol Biol.* 407 (2), 193–221.

Turan, S; Zehe, C; Kuehle, J; Qiao, J; Bode, J (2013). Recombinase-mediated cassette exchange (RMCE) - a rapidly-expanding toolbox for targeted genomic modifications. *Gene* 515 (1), 1–27.

Val, ME; Marbouty, M; de Lemos Martins, F; Kennedy, SP.; Kemble, H; Bland, MJ. *et al*., (2016). A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci Adv.* 2 (4), e1501914.

van Dijk, EL; Auger, H; Jaszczyszyn, Y; Thermes, C (2014). Ten years of next-generation sequencing technology. *Trends Genet.* 30 (9), 418–426.

Van Duyne, GD (2001). A structural view of *cre-loxp* site-specific recombination. *Annu Rev Biophys Biomol Struct.* 30, 87–104.

Vasu, K; Nagamalleswari, E; Nagaraja, V (2012). Promiscuous restriction is a cellular defense strategy that confers fitness advantage to bacteria. *Proc Natl Acad Sci U S A* 109 (20), E1287-93.

Waaijers, S; Boxem, M (2014). Engineering the *Caenorhabditis elegans* genome with CRISPR/Cas9. *Methods* 68 (3), 381–388.

Wang, B; Kitney, RI; Joly, N; Buck, M (2011a). Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat Commun.* 2, 508.

Wang, Y; Yau, YY; Perkins-Balding, D; Thomson, JG (2011b). Recombinase technology: applications and possibilities. *Plant Cell Rep.* 30 (3), 267-85.

Wang, X; Rudner, DZ (2014). Spatial organization of bacterial chromosomes. *Curr Opin Microbiol.* 22, 66–72.

Wang, Y; Zhang, ZT; Seo, SO; Choi, K; Lu, T; Jin, YS; Blaschek, HP (2015). Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9 system. *J Biotechnol.* 200, 1–5.

Wang, Y; Cobb, RE; Zhao, H (2016a). High-Efficiency Genome Editing of *Streptomyces* Species by an Engineered CRISPR/Cas System. *Methods Enzymol.*575, 271–284.

Wang, Y; Zhang, ZT; Seo, SO; Lynn, P; Lu, T; Jin, YS; Blaschek, HP (2016b). Bacterial Genome Editing with CRISPR-Cas9: Deletion, Integration, Single Nucleotide Modification, and Desirable "Clean" Mutant Selection in *Clostridium beijerinckii* as an Example. *ACS Synth Biol.* 5 (7), 721–732.

Weil, J; Signer, ER (1968). Recombination in bacteriophage lambda. II. Site-specific recombination promoted by the integration system. *J Mol Biol.*34 (2), 273–279.

Werner, S; Engler, C; Weber, E; Gruetzner, R; Marillonnet, S (2012). Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioeng Bugs* 3 (1), 38–43.

Westbrook, AW; Moo-Young, M; Chou, CP (2016). Development of a CRISPR-Cas9 Tool Kit for Comprehensive Engineering of *Bacillus subtilis*. *Appl Environ Microbiol.* 82 (16), 4876–4895.

Wiktor, J; Lesterlin, C; Sherratt, DJ; Dekker, C (2016). CRISPR-mediated control of the bacterial initiation of replication. *Nucleic Acids Res.* 44 (8), 3801–3810.

Wilson, JH; Berget, PB; Pipas, JM (1982). Somatic cells efficiently join unrelated DNA segments end-to-end. *Mol Cell Biol.* 2 (10), 1258–1269.

Wirth, D; Gama-Norton, L; Riemer, P; Sandhu, U; Schucht, R; Hauser, H (2007). Road to precision: recombinase-based targeting technologies for genome engineering. *Curr Opin Biotechnol*. 18 (5), 411–419.

Withers, HL; Bernander, R (1998). Characterization of *dnaC2* and *dnaC28* mutants by flow cytometry. *J Bacteriol.* 180 (7), 1624–1631.

Wright, R; Stephens, C; Shapiro, L (1997). The CcrM DNA methyltransferase is widespread in the alpha subdivision of proteobacteria, and its essential functions are conserved in *Rhizobium meliloti* and *Caulobacter crescentus*. *J Bacteriol*. 179 (18), 5869–5877.

Wu, G; Yan, Q; Jones, JA; Tang, YJ; Fong, SS; Koffas, MA (2016). Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol.* 34 (8), 652–664.

Xie, T; Fu, LY; Yang, QY; Xiong, H; Xu, H; Ma, BG; Zhang, HY (2015). Spatial features for *Escherichia coli* genome organization. *BMC Genomics* 16, 37.

Xue, GP; Johnson, JS; Dalrymple, BP (1999). High osmolarity improves the electro-transformation efficiency of the gram-positive bacteria *Bacillus subtilis* and *Bacillus licheniformis*. *J. Microbiol. Methods* 34 (3), 183–191.

Xue, X; Wang, T; Jiang, P; Shao, Y; Zhou, M; Zhong, L *et al*., (2015). MEGA (Multiple Essential Genes Assembling) deletion and replacement method for genome reduction in *Escherichia coli*. *ACS Synth Biol.* 4 (6), 700–706.

Yeeles, JT; Dillingham, MS (2010). The processing of double-stranded DNA breaks for recombinational repair by helicase-nuclease complexes. *DNA Repair* 9 (3), 276–285.

Yip, CB; Ding, H; Hynes, MF (2015). Counter-transcribed RNAs of *Rhizobium leguminosarum repABC* plasmids exert incompatibility effects only when highly expressed. *Plasmid* 78, 37–47.

Yonemura, I; Nakada, K; Sato, A; Hayashi, JI; Fujita, K; Kaneko, S; Itaya, M (2007). Direct cloning of full-length mouse mitochondrial DNA using a *Bacillus subtilis* genome vector. *Gene* 391 (1-2), 171–177.

Zhang, K; Duan, X; Wu, J (2016). Multigene disruption in undomesticated *Bacillus subtilis* ATCC 6051a using the CRISPR/Cas9 system. *Sci Rep*. 6, 27943.

Zhou, J; Wu, R; Xue, X; Qin, Z (2016). CasHRA (Cas9-facilitated Homologous Recombination Assembly) method of constructing megabase-sized DNA. *Nucleic Acids Res*. 44 (14), e124.

Zuñiga-Castillo, J; Romero, D; Martínez-Salazar, JM (2004). The recombination genes *addAB* are not restricted to gram-positive bacteria: genetic analysis of the recombination initiation enzymes RecF and AddAB in *Rhizobium etli*. *J Bacteriol.* 186 (23), 7905–7913.

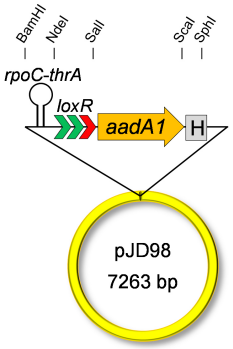
6. Anhang

6.1 Konstruktion der *S. meliloti* Fusionsstämme SmAB und SmABC

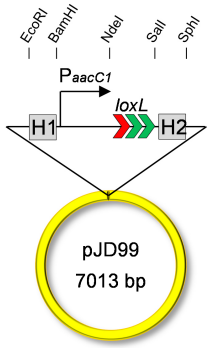
Zur Herstellung der Megaplasmid-Fusion pSymAB wurde pJD99 mittels Konjugation stromaufwärts von *nodP1* auf pSymA in *S. melilioti* JDSm57 (Döhlemann *et al*., 2016) integriert. Nach Saccharose-Selektion vermittelter Deletion des Plasmid-*Backbones* wurde pJD98 mittels einfacher HR stromabwärts von *nodQ2* auf pSymB integriert. Die Taurin-induzierte Expression des Rekombinasegens *cre* führte zur Fusion der Megaplasmide. Die spezifische Anordnung der funktionalen Elemente stromauf- bzw. abwärts der homologen *nodPQ*-Cluster ermöglichte die Saccharose-Selektion vermittelte Deletion des B*ackbones* von pJD98, einer funktionalen *loxP* Sequenz sowie des Transkriptionsterminators. *S. meliloti* SmAB weist aus diesem Grund lediglich eine Kopie des *nodPQ*-Clusters auf.
Die anschließende Konstruktion von SmABC erfolgte durch Integration von pJD130 stromaufwärts von *algI* auf pSymAB via Konjugation. Nachdem das Plasmid-*Backbone* mittels Saccharose-Selektion deletiert wurde, konnte pJD126 stromabwäts eines *algI*-Homologs auf dem Chromosom integriert werden. Eine Cre/*lox*-vermittelte Rekombination führte zur vollständigen Fusion der Replikons. Die Klonierungsstrategie erlaubte schließlich, analog zur Herstellung von SmAB, die Saccharose-Selektion vermittelte Deletion des *Backbones* von pJD126 sowie funktioneller Elemente, wodurch anschließende genetische Manipulationen auf Basis von pK18mob-Derivaten ermöglicht wurden.

Plasmid-Konstruktion von pJD98, pJD99, pJD126 und pJD130:

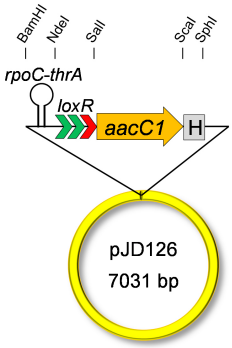
pJD98 ist ein Derivat von pK18mobsacB mit vier *Inserts*. Der Transkriptionsterminator *thrA-rpoC* wurde mit den Primern 131 (5‘-cgcggatccctgacgcgtacaggaaacac) und 132 (5‘-gcgcatatgcgggctctttccctaaactcc) von genomischer DNA aus *S. meliloti* B033 (Harrison *et al*., 2011) amplifiziert. Eine *loxR* Sequenz wurde mittels Hybridisierung der Primer 189 (5‘ tatgataacttcgtatagcatacatt atacgaacggtag) und 190 (5‘-tcgactaccgttcgtataatgtatg ctatacgaagttatca) hergestellt. Das Spectinomycin-Resistenzgen *aadA1* wurde von pJG181 (Harrison *et al*., 2011) mittels der Primer 127 (5‘-cgcgtcgac ggactcaatacaccatgaggg) und 128 (5‘-gcgctgcagagtactttatttgccgactac cttggtg) PCR-amplifiziert, während ein 500 bp großes Fragment (H) zur homologen Rekombination nahe *nodPQ2* mit den Primern 240 (5’-cgcccct ttgcgaaatgct) und 241 (5‘-cgcgcatgctggcaaccgcggtcagttc) von genomischer DNA aus *S. meliloti* Rm1021 amplifiziert wurde.



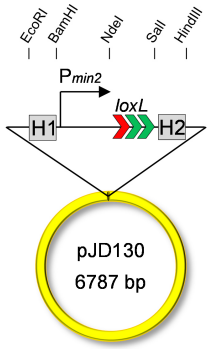
pJD99 ist ein pK18mobsacB-Derivat mit vier *Inserts*. ~500 bp große homologe Regionen zur Integration nahe *nodPQ* wurden mit den Primern 236 (5‘-cgcgaattctggagaccgggtcaagctg) und 237 (5‘-cgcggatcccgacttcggtaccggtgcc) (H1), sowie 238 (5‘-gcggtcgaccgtccgccgttgcgtatctc) und 239 (5’-cgcgcat gcctatccgggttacgcctcgg) (H2) von genomischer DNA aus *S. meliloti* Rm1021 amplifiziert. Der ~300 bp große Gentamicin-Resistenzkassettenpromotor P*aacC1* wurde von pMS252 (Becker *et al*., 1995) amplifiziert. Eine *loxL* Sequenz wurde mittels Hybridisierung der Primer 187 (5’-tatgtaccgttcgt atagcatacattatacgaagttatg) und 188 (5’-tcgacataacttcgtataatgtatgctatacgaac ggtaca) hergestellt.



pJD126 ist ein Derivat von pK18mobsacB mit vier *Inserts*. Der Transkriptionsterminator *thrA-rpoC* wurde mit den Primern 131 (5'-cgcggatc cctgacgcgtacaggaaacac) und 132 (5'-gcgcatatgcgggctctttccctaaactcc) von genomischer DNA aus *S. meliloti* B033 (Harrison *et al.*, 2011) amplifiziert. Eine *loxR* Sequenz wurde mittels Hybridisierung der Primer 189 (5' tatgataacttcgtatagcatacatt atacgaacggtag) und 190 (5'-tcgactaccgttcgtataatgta tgctatacgaagttatca) hergestellt. Das Gentamicin-Resistenzgen *aacC1* wurde von pJG181 (Harrison *et al.*, 2011) mittels der Primer 125 (5'-gcggtcgacgttat ggagcagcaacgatgtt) und 126 (5' gcgctgc agagtactttaggtggc ggtacttgggtcg) PCR-amplifiziert, während ein 500 bp großes Fragment (H) zur homologen Rekombination nahe *algI* mit den Primern 364 (5'-gaaagtggattccgagtcttc) und 365 (5'-tccaatctccgtttgtcgttg) von gen. DNA aus *S. meliloti* Rm1021 amplifiziert wurde.



pJD130 ist ein pK18mobsacB-Derivat mit vier *Inserts*. ~500 bp große homologe Regionen zur Integration nahe eines *algI*-Homologs wurden mit den Primern 360 (5'-cgcgaattctgaaaatagcggtatagcgatgtggc) und 361 (5'-gcgggatcccc gcttgcgcgtcttctcg) (H1), sowie 362 (5'-cgcctgcaggtgctggtcgccggccagg) und 363 (5'-gcgaagcttgcggatcgaaagtgataactgccatag) (H2) amplifiziert. Die 55 bp kleine P*min2*-Promotorregion (Δ *Shine-Dalgarno*) wurde mittels Hybridisierung der Primer 388 (5'-gatcctgttgacaattaatcatcgaactagttaactagtacgcaagtagacgcgttacgtca) und 389 (5'-tatgacgtaacgcgtctacttgcgtactagttaactagttcgatgattaattgtcaacag) generiert, während zur Herstellung der *loxL* Sequenz die Primer 187 (5'-tatgtaccgttcgtatagcatacattatacgaagttatg) und 188 (5'-tcgacataacttcgtataatgtatg ctatacgaacggtaca) hybridisiert wurden.



6.2 Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle

Eigener Beitrag zur Studie:

I. Konstruktion des artifiziellen Mini-Replikons pART

II. Herstellung der *S. meliloti* Stämme Sma818 DnaN-mCh, pART, pART ParB-mCherry, 2011 pSymAter-eYFP und 2011 pSymBter-eYFP

III. Fluoreszenzmikroskopische Analysen der in II. genannten Stämme in Zusammenarbeit mit BF

IV. Probenvorbereitung zur *marker frequency analysis*

(BF, Benjamin Frage)

Molecular Microbiology (2016) 100(5), 808–823 ■
doi:10.1111/mmi.13351
First published online 20 April 2016

Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle

Benjamin Frage,[1] Johannes Döhlemann,[1] Marta Robledo,[1†] Daniella Lucena,[2] Patrick Sobetzko,[1] Peter L. Graumann[2] and Anke Becker[1*]

[1]*LOEWE Center for Synthetic Microbiology and Faculty of Biology, Philipps-Universität Marburg, 35032, Marburg, Germany.*

[2]*LOEWE Center for Synthetic Microbiology and Faculty of Chemistry, Philipps-Universität Marburg, 35032, Marburg, Germany.*

Summary

A considerable share of bacterial species maintains multipartite genomes. Precise coordination of genome replication and segregation with cell growth and division is vital for proliferation of these bacteria. The α-proteobacterium *Sinorhizobium meliloti* possesses a tripartite genome composed of one chromosome and the megaplasmids pSymA and pSymB. Here, we investigated the spatiotemporal pattern of segregation of these *S. meliloti* replicons at single cell level. Duplication of chromosomal and megaplasmid origins of replication occurred spatially and temporally separated, and only once per cell cycle. Tracking of FROS (fluorescent repressor operator system)-labelled origins revealed a strict temporal order of segregation events commencing with the chromosome followed by pSymA and then by pSymB. The *repA2B2C2* region derived from pSymA was sufficient to confer the spatiotemporal behaviour of this megaplasmid to a small plasmid. Altering activity of the ubiquitous prokaryotic replication initiator DnaA, either positively or negatively, resulted in an increase in replication initiation events or G1 arrest of the chromosome only. This suggests that interference with DnaA activity does not affect replication initiation control of the megaplasmids.

Accepted 5 February, 2016. *For correspondence. E-mail anke.becker@synmikro.uni-marburg.de; Tel. (+49) 6421 28 24451; Fax (+49) 6421 28 22229. †Present address: Grupo de Ecología Genética de la Rizosfera, Estación Experimental del Zaidín (CSIC), Granada, Spain

Introduction

Faithful maintenance of genome content is crucial to ensure bacterial proliferation. Hence, failure to reliably coordinate DNA replication and segregation with cell growth and division is deleterious to survival. Bacterial genome architectures are relatively uniform with the majority of sequenced genomes consisting of a single circular chromosome. However, about 10% of the sequenced bacterial species maintain multipartite genomes. These comprise diverse phyla ranging from Chloroflexi, Deinococcus – Thermus, Spirochaetes to α-, β- and γ-proteobacteria (Val *et al.*, 2014). To date, the best studied examples are γ-proteobacteria of the Vibrionaceae family carrying two chromosomes of uneven sizes, with the secondary smaller chromosome being attributed to ancestral plasmid domestication (Kirkup *et al.*, 2010; Okada *et al.*, 2005). Further prominent examples of secondary large replicons are RepABC family megaplasmids and secondary chromosomes exclusively found in α-proteobacteria (Pinto *et al.*, 2012). Amongst these are plant-pathogenic *Agrobacterium* and nitrogen-fixing plant-symbiotic α-rhizobial, as well as human- and animal-pathogenic *Brucella* and *Bartonella* species. *Rhizobium* species contain up to six medium to large size RepABC family replicons in addition to the primary chromosome (González *et al.*, 2006).

In contrast to plasmids initiating replication multiple times during the course of the cell cycle (del Solar *et al.*, 1998), chromosomal replication is usually initiated only once per generation at a fixed time point (Katayama *et al.*, 2010). However, this distinction is not true for megaplasmids and secondary chromosomes which alike the primary chromosome replicate only once per generation (Pinto *et al.*, 2012; Val *et al.*, 2014). This raises questions if and how replication initiation and segregation of multiple replicons is coordinated in time and space in bacterial cells.

Cell cycle control circuits of the asymmetrically dividing α-proteobacteria *Caulobacter crescentus* and *Sinorhizobium meliloti* show striking similarities (Lam *et al.*, 2003; Hallez *et al.*, 2004; Fields *et al.*, 2012; Pini *et al.*, 2013; De Nisco *et al.*, 2014; Pini *et al.*, 2015; Panis *et al.*, 2015). Whilst *C. crescentus* possesses a single chromosome (4.02 Mb) (Nierman *et al.*, 2001), the *S.*

meliloti genome is composed of a main chromosome (3.65 Mb) and the megaplamids pSymA (1.35 Mb) and pSymB (1.68 Mb) (Galibert *et al*. 2001). Like almost all bacteria, these two α-proteobacteria rely on the highly conserved DnaA for commencing replication of their main chromosome (Mott and Berger, 2007; Zweiger and Shapiro, 1994; Sibley *et al*., 2006). DnaA, a member of the AAA+ protein family, binds ATP-dependently to a set of sites within the chromosomal origin of replication (*Cori*), forming an initiation complex that unwinds the AT-rich region of the origin and recruits the replication machinery (Kaguni, 2006; Katayama *et al*., 2010). Tight control of DnaA is a vital component of negative feedback regulation linked to genome replication. Studies in *Escherichia coli*, *C. crescentus* and *Bacillus subtilis*, revealed multiple regulatory mechanisms exerted at different levels of control encompassing repression of *dnaA* transcription and sequestration of *Cori*, as well as proteolysis, titration and regulatory inactivation of DnaA (Wang and Kaguni, 1987; Lu *et al*., 1994; Soufo *et al*., 2008; Katayama *et al*., 1996; Jonas *et al*., 2013; Kitagawa *et al*., 1998; Kato and Katayama, 2001).

As in *E. coli*, *R*egulatory *I*nactivation of *D*na*A* (RIDA) is an important factor of replication initiation control in *C. crescentus* (Collier and Shapiro, 2009; Jonas *et al*., 2011) suggesting that this mechanism may also be conserved in *S. meliloti*. RIDA promotes hydrolysis of ATP bound to DnaA resulting in DnaA-ADP, which is inactive for initiation (Kato and Katayama, 2001). It is catalysed by a nucleoprotein complex consisting of duplex DNA, the DNA-loaded beta clamp DnaN, and the ADP-bound Hda protein (Su'etsugu *et al*., 2008). Both the *E. coli* Hda and the *C. crescentus* homologue HdaA share similarities with the AAA+ ATPase domain of their cognate DnaA protein (Kato and Katayama 2001; Collier and Shapiro 2009).

The *S. meliloti* megaplasmids pSymA and pSymB belong to the RepABC family that is characterised by the combined replication and partitioning *repABC* locus (Cervantes-Rivera *et al*., 2011; Pinto *et al*. 2012). Whilst RepC most likely acts as the replication initiator protein at the origin of replication, RepA, RepB and the partitioning sites are required for vertical transmission of the megaplasmids to daughter cells.

Bacterial genomes exhibit distinct spatial orientations in the cell. In *C. crescentus*, the replication origin and terminus of the circular chromosome reside near the old and new cell pole respectively (Viollier *et al*., 2004). Snapshot studies of *S. meliloti* cells showed the replication origins of the chromosome and both megaplasmids (*Cori*, *SymAori* and *SymBori*) also preferably localised to the poles, with some drift away from the poles of *SymAori* and *SymBori* (Kahng and Shapiro, 2003).

So far, *Vibrio cholerae* is the only well-studied model for multiple chromosome maintenance and replication. In this γ-proteobacterium, a delay in replication initiation of the smaller chromosome 2 (1.07 Mb) leads to termination synchrony with the larger chromosome 1 (2.96 Mb) (Rasmussen *et al*., 2007). In contrast, in the class of α-proteobacteria multipartite genome replication and partitioning have been poorly investigated and to date very little is known about timing and spatial organisation of these processes in bacteria possessing genomes divided in more than two replicons. Recently, the *Brucella abortus* chromosomes I (2.1 Mb) and II (1.2 Mb) have been found to be oriented along the cell length axis, similar to the chromosomal organisation of *C. crescentus*, and the origin of chromosome II to be duplicated after that of chromosome I (Deghelt *et al*., 2014).

Here we report on the spatiotemporal choreography of the *S. meliloti* tripartite genome during the cell cycle. We show that partitioning of the three replicons follow a highly ordered succession and that the *repABC* region alone is sufficient to maintain the native spatiotemporal behaviour of a secondary RepABC family replicon. Evidence is presented for alterations in DnaA activity affecting the abundance of the main chromosome only. We speculate that the distinct timings of megaplasmid segregation rely on spatial checkpoints involving genetic components located within the *repABC* cassette.

Results

Intracellular localisation and dynamics of the DNA polymerase III beta subunit DnaN in S. meliloti

To determine the initiation pattern of the *S. meliloti* replication machinery, a fusion of the fluorescent reporter mCherry to the DNA polymerase III beta subunit DnaN was employed to visualise intracellular occurrence and dynamics of the replisome in a time lapse at 2 min intervals. The gene fusion replaced the *dnaN* wild-type allele in single copy in the chromosome under the control of its native promoter. This replacement had no effect on cell morphology or growth implying full functionality of the DnaN-mCherry fusion as previously described for similar reporter constructs in other bacteria (Collier and Shapiro, 2009; Veening *et al*., 2009).

Throughout the manuscript, time point zero marks physical separation of sibling cells. Immediately after cytokinesis, DnaN-mCherry localised to the old cell pole of one daughter cell indicating direct entry into the S phase state (hereafter called S cell), whilst in the other daughter cell the fluorescent signal was first dispersed indicating arrest in the G1 phase until DnaN-mCherry also localised to the old pole (hereafter called G1 cell) (Fig. 1A). Completion of the cell cycle of the S and G1

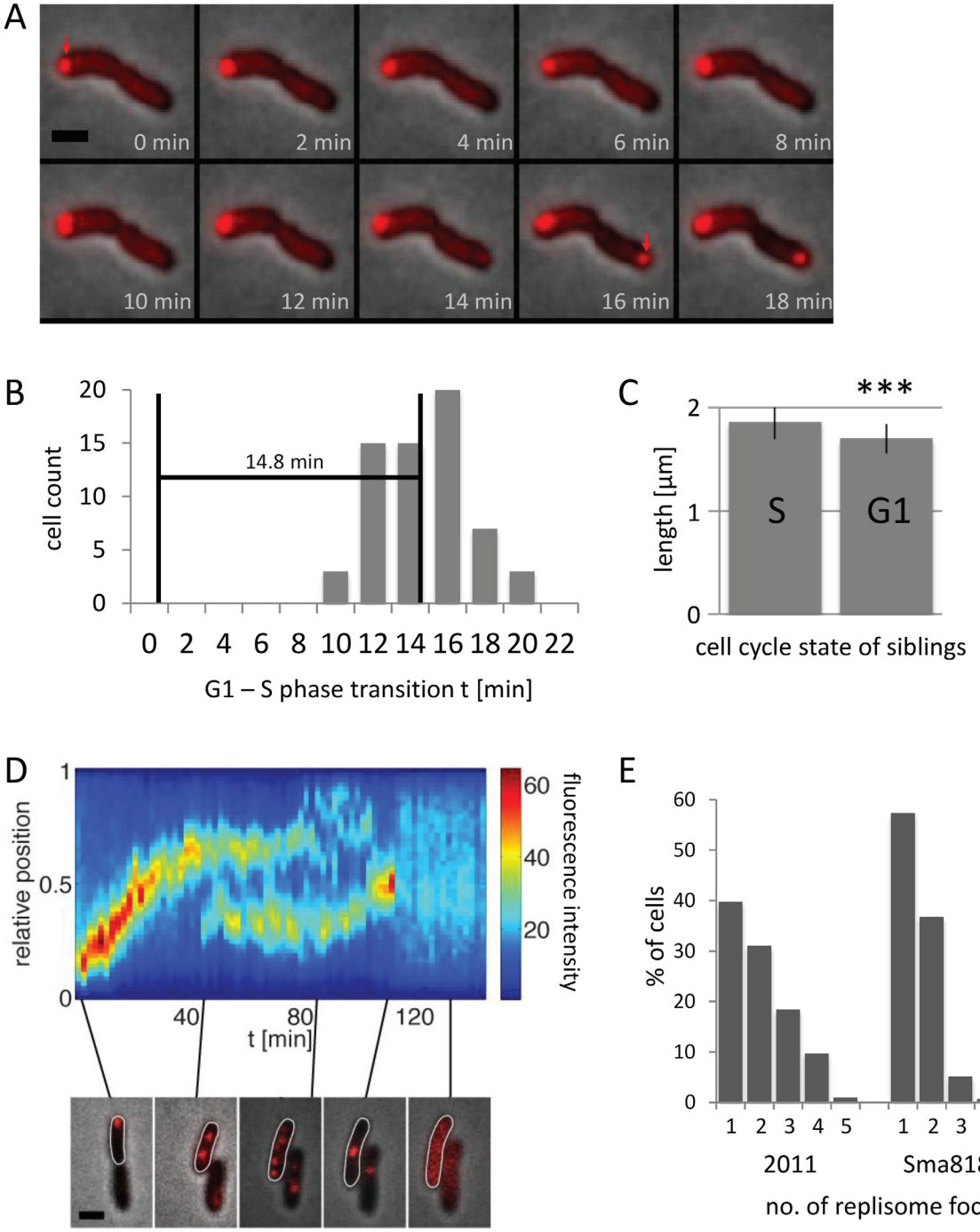


Fig. 1. Spatiotemporal dynamics of the *S. meliloti* replisome.
A. Time-lapse fluorescence microscopy analysis of dividing cells expressing *dnaN-mCherry* displays asymmetric timing of replication initiation. First appearances of replisome clusters are labelled with arrows. Two minute time points, scale bar 2 µm.
B. Time span of smaller daughter cells staying in G1 phase. Sixty three dividing cells were analysed.
C. Newborn cells with a polar DnaN-mCherry focus are statistically longer than their siblings (***$P < 0.0001$). Length correlation of 48 newborn cells with a polar DnaN-mCherry focus (S phase cell) to that of their 48 siblings without a polar focus (G1 phase cell) is shown.
D. Representative kymograph of replisome dynamics. DnaN-mCherry fluorescence intensities determined at 2 min intervals were mapped to their relative positions within the cell (0 = old pole, 1 = new pole). Microscopy snapshots of characteristic replisome dynamics are shown for (t = 0, t = 45, t = 85, t = 110, t = 130 min), scale bar 2 µm.
E. Abundance of cells with different numbers of DnaN-mCherry foci in 2011 (wild-type) and Sma818 (delta pSymA) strains. $n > 100$ cells for each strain.

cell took 141.2 ± 7.5 and 157.5 ± 8.6 min ($n = 16$) respectively and a 14.8 ± 5.0 min delay was determined until the G1 cell entered the S phase (Fig. 1B).

Several α-proteobacteria including *S. meliloti* divide asymmetrically resulting in daughter cells of different lengths (Lam *et al*., 2003; Hallez *et al*., 2004). We determined a ratio of 1.10 between daughter cells, with one cell being about 10% longer (1.87 ± 0.17 μm) than the other (1.7 ± 0.15 μm). S cells, indicated by immediate polar DnaN-mCherry localisation after division, were statistically longer than their G1 siblings. A paired *t*-test of daughter cell lengths with polar-localised vs. dispersed DnaN-mCherry signal gave a two-tailed *P* value of less than 0.0001, which can be considered as statistically highly significant upon conventional criteria (Fig. 1C).

During the further course of the cell cycle, the polarly localised DnaN-mCherry spot gradually moved towards the new pole to a position corresponding to about two-thirds of the cell length (Fig. 1D and Supporting Information Video S1). 37.7 ± 4.8 min ($n = 20$) after appearance of this first DnaN-mCherry spot, a second distinct focus emerged at about one-third distance of the cell length from the old pole (Fig. 1D and Supporting Information Video S1). At this stage, kymographs of representative time-lapse data generally showed two main clusters that remained spatially separated (Fig. 1D). Later in the cell cycle, more than two foci or large DnaN clusters showing a highly dynamic localisation pattern, and rarely up to five distinct foci, were observed (Fig. 1D and Supporting Information Video S1). Towards the end of the cell cycle, before DnaN-mCherry became dispersed, the number of foci dropped back to one (Fig. 1D and Video S1). Upon disappearance of the last DnaN focus, cells took another 47.4 ± 9.9 min ($n = 20$) before cell division events were observed. Kymograph analysis of the shorter daughter cells indicates similar replisome dynamics, with exception of the delay of about 15 min before DnaN-mCherry accumulated in the old pole region (Supporting Information Fig. S1A).

The multiple DnaN-mCherry foci probably are indicative of multiple spatially separated replisomes of the chromosome and/or the megaplasmids. Marker frequency analysis of logarithmic vs. stationary phase *S. meliloti* cultures confirmed bidirectional replication starting from a single origin of each of the three replicons (Supporting Information Fig. S1B). This supports the assumption that the additional DnaN foci are associated to replication of the megaplasmids. We introduced the chromosomal *dnaN-mCherry* fusion to *S. meliloti* Sma818, a strain cured of the pSymA megaplasmid (Oresnik et al., 2000). When comparing this strain to the wild-type, more pSymA-cured cells displayed one or two DnaN-mCherry foci, whereas less cells showed three or four foci and none more than four foci (Fig. 1E). This further supports the assumption that multiple additional DnaN foci are associated to megaplasmid replication.

Segregation of the three S. meliloti *replicons is a temporally and spatially ordered process*

To gain insights into the segregation dynamics of the three *S. meliloti* replicons, we performed origin-tracking experiments applying the fluorescent repressor operator system (FROS) for observation of chromosomal, pSymA and pSymB origins. Following cell division, chromosomal origins segregated after 18 ± 4 min in the S cell (cell cycle of 139 ± 12 min), whilst in the G1 cell segregation was observed after 32 ± 8 min (cell cycle of 154 ± 14 min) (Fig. 2A, B and Video S2). The megaplasmid origins (*SymAori*-eYFP, *SymBori*-eCFP) showed different segregation timings. *SymAori* segregation occurred post division at 43 ± 6 min in S cells and at 62 ± 4 min in G1 cells. *SymBori* segregation was even further delayed, with 56 ± 8 min in S cells and 68 ± 6 min in G1 cells (see arrows Fig. 2A and Supporting Information Videos S3, S4).

Tracking of the three *S. meliloti* origins, and kymographs of representative individual time-lapse experiments illustrate the spatiotemporal pattern of origin segregation (Figs 2B and Supporting Information Fig. S2A). Similar to *Cori* dynamics observed in *C. crescentus* (Shebelut *et al*., 2010), dynamics of the three types of *S. meliloti* origins share a basic pattern that can be separated into three distinct phases: (i) the release of a single origin focus from the old pole, (ii) the appearance of the second focus and retraction of one of these origins to the old pole and (iii) concurrent translocation of the other origin to the new pole. Upon polar release (i), a single distinct origin focus per replicon moved in a linear fashion towards midcell. *Cori* moved 0.1–0.3 ($n = 10$) of the cell length away from the old pole before a second *Cori* spot appeared after 17.5 ± 4.1 min (Fig. 2B and C, see arrows Supporting Information Fig. S2A). Upon appearance of a second *Cori* signal, translocation of one of these foci to the new pole occurred at higher speed than during the release phase and was completed within 20.3 ± 3.8 min ($n = 50$). Before segregation events of the megaplasmids were observed, single pSymA and pSymB origins continued to move towards midcell, implying an extended polar release phase (i). Although the average trajectory within phase (i) was linear, it was slower and origins underwent higher spatial variation between the 5 min monitoring intervals than in later phases, including receding movements towards the old pole at random time points (Supporting Information Fig. S2B, see

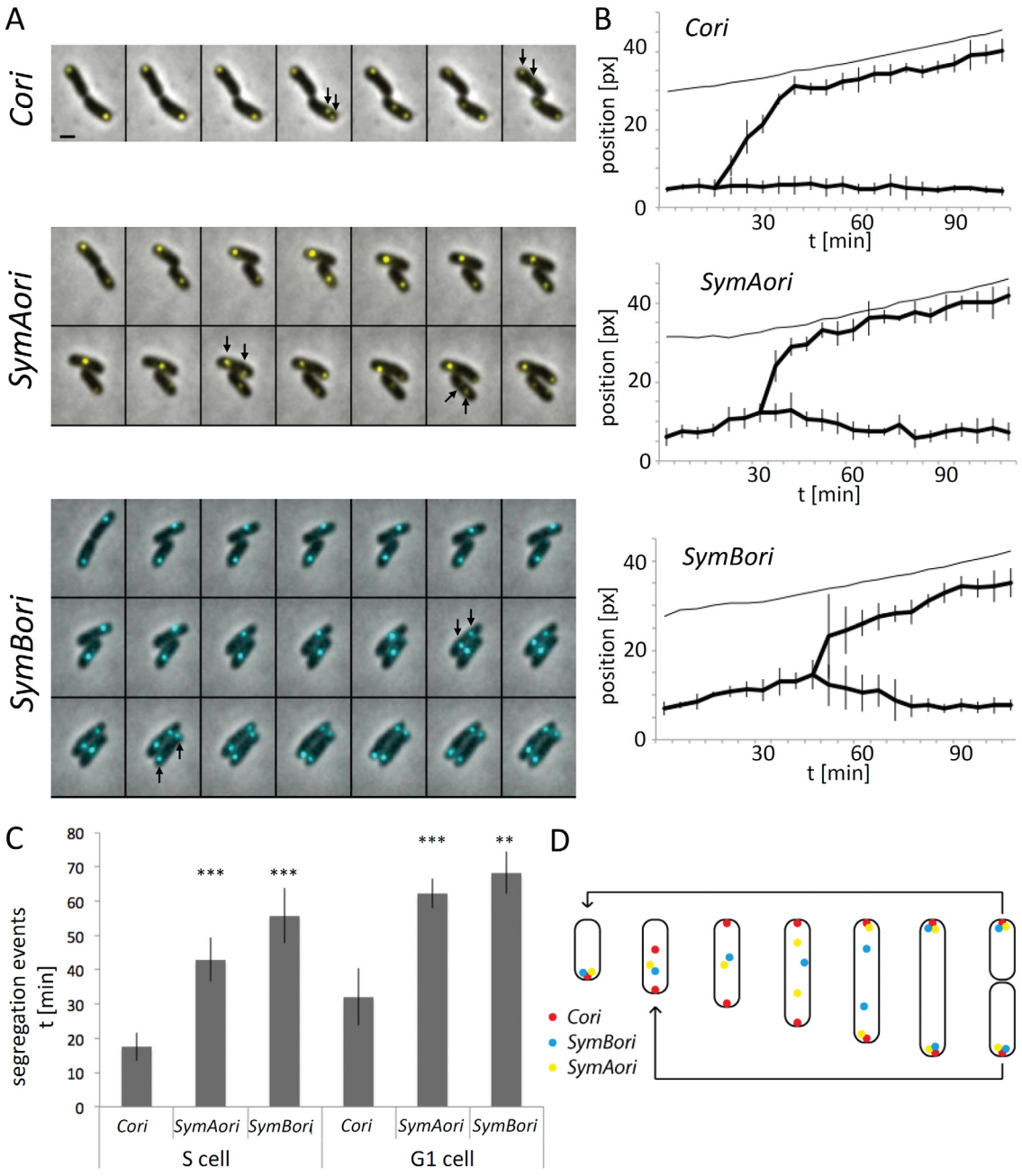


Fig. 2. Spatiotemporal dynamics of replicon origins.
A. Representative time-lapse fluorescence microscopy images of cells carrying carrying *Cori*-eYFP, *SymAori*-eYFP, or *SymBori*-eCFP FROS-tagged origins. Shown are overlays of fluorescence and phase contrast channels of dividing cells until origin segregation events were observed (see arrows). Five minute time points, scale bar 1 μm.
B. Tracking of *Cori*, *SymAori* or *SymBori* origins. Upper thin lines display cell lengths. Thick lines represent medians ($n = 4$) of relative origin positions within the cell at every 5 min over a whole division cycle. Position [px] 0 = old pole. C. Statistical significant (***$P < 0.001$; **$P < 0.005$) temporal order of replicon segregation events after cell division, divided into S and G1 phase daughter cells. Chromosome: S cell 17.5 ± 4.1 min, G1 cell 32 ± 8.3 min, $n = 40$. pSymA: S cell 42.9 ± 6.4 min, G1 cell 62.1 ± 4.3 min, $n = 28$. pSymB: S cell 55.7 ± 8.1 min, G1 cell 68.2 ± 6.1 min, $n = 28$.
D. Schematic model of *S. meliloti* origin dynamics and segregation events.

arrows). Paired *t*-tests comparing segregation timings of *Cori-SymAori* or *SymAori-SymBori* in S cells, as well as *Cori-SymAori* in G1 cells resulted in *P* values < 0.001, which can be considered as statistically highly significant. A paired *t*-test of G1 cell *SymAori-SymBori* data gave a significant *P* value of < 0.005. Whilst the timing of replicon segregation was strictly ordered, initial positions of origin segregation relative to the old pole and

cell length ranged from 0.3 to 0.5 ($n = 10$) for *SymAori*, and 0.4 to 0.7 ($n = 10$) for *SymBori* (Figs 2B and Supporting Information Fig. S2B). Time-lapse studies of *SymBori*-eCFP (Fig. 2A) and *SymBori*-eYFP FROS labels (Supporting Information Fig. S3A and Videos S4, S5) led to similar results. Plotting of the origin tracks of dual-color FROS visualisations of *Cori* and *SymBori*, or *SymAori* and *SymBori* revealed another interesting feature (Fig. 3A and B). Whilst *Cori*s moved to the extreme ends of the cell, in the same cell *SymBori*s localised near the poles after completion of their segregation, but never reached the extreme positions of the chromosomal origins. *SymAori* showed trajectories similar to *SymBori* and localisation of the associated signals nearby the poles overlapped with those of *SymBori* (Fig. 3B).

Simultaneous tracking of *Cori* and *SymBori*, or *SymAori* and *SymBori* in individual cells further confirmed the delay between segregation events, with *SymBori* segregating 35.4 ± 5.8 min after *Cori*, and *SymAori* segregating 20.4 ± 4 min before *SymBori* (Fig. 3A and Supporting Information Videos S6, S7). Furthermore, *SymBori* segregation did not commence before *Cori* segregation was complete, indicated by arrival of these origins at the cell poles. Although the destination positions of *SymAori*s in the pole regions varied, it seems that partitioning of *SymBori*s was not initiated before segregation of *SymAori*s was complete. The translocation phase of one of the *Coris* to the new pole never took longer than 20 min, after the 18 min polar release phase. In this release phase, cells were already in S phase, as indicated by formation of the polar DnaN focus (Fig. 1), but without segregation being observed (Fig. 2). Therefore, partitioning of the chromosomal origins likely was complete, before pSymA segregation started at 43 min (Fig. 2C). Translocation of *SymAori*s from around midcell was finished after 10–15 min and simultaneous observation of the origins of replication of both megaplasmids never showed segregation of the *SymBori*s before the *SymAori*s arrived at their near-pole positions (Fig. 3B). Figure 2D summarises the strict temporal order of origin duplication and segregation events with the chromosome coming first, followed by pSymA and pSymB. The *dnaN-mCherry* construct was introduced into a strain carrying the *Cori*-eYFP and *SymBori*-eCFP FROS labels. Additional DnaN-mCherry foci were observed after *SymBori* segregation had become visible (Supporting Information Fig. S3B, see arrows; Video S8), further supporting the assumption that replication initiation of the megaplasmids results in additional DnaN foci.

To determine the spatiotemporal pattern of the terminus regions of both megaplasmids SymAter-eYFP and SymBter-eYFP FROS labels were monitored. The terminus region of the smaller pSymA segregated at 70.5 ± 5.9 min (S cells, cell cycle of 124.5 ± 9.3 min) and 92.8 ± 10.3 min (G1 cells, cell cycle of 155.5 ± 15.3 min) ($n = 20$), whilst the terminus region of pSymB segregated at 93.5 ± 6.7 min (S cells, cell cycle of 130.0 ± 6.7 min) and 111.9 ± 7.0 min (G1 cells, cell cycle 155.0 ± 13.5 min) ($n = 20$) after cell division (Fig. 4 and Supporting Information Videos S9, S10). Consistent generation times in the experiments monitoring the origins of replication (Fig. 2) and the terminus regions (Fig. 4) allow for deducing the probable interval between segregation of origin and terminus of each megaplasmid. We deduced an interval of 30 min for pSymA and of about 40 min for pSymB, which is in good agreement with the difference in size of these megaplasmids. The terminus region of pSymA localised around midcell, and after segregation at approximately 1/4 and 3/4 positions of the predivisional cell (Fig. 4 and Supporting Information Video S9). The terminus region of pSymB first localised near the new pole before localisation to midcell was observed (Fig. 4 and Supporting Information Video S10).

The RepABC cassette alone is sufficient for segregation timing

Origin-tracking data showed that segregation of the smaller pSymA starts earlier than that of pSymB, excluding that the order of these events correlates with replicon sizes. Applying the FROS system, we examined the dynamics of the 11.3 kb mini-replicon pART consisting basically of the pSymA *repA2B2C2* region and a 4.1 kb *tetO* array, introduced to *S. meliloti* Sma818, a strain that was cured of its pSymA replicon. The spatiotemporal patterns of segregation of pART and its 1.35 Mb FROS-tagged ancestor pSymA were comparable. As observed for pSymA, after cell division the origin of pART was released from the old pole and moved towards midcell followed by segregation of the origins, which started at about 47.5 ± 5.4 min in the S phase cell, or at 55.0 ± 6.7 min in the G1 cell (Fig. 3C, D and Supporting Information Video S11). A paired *t*-test of pART segregation events in S and G1 cells gave a *P* value of < 0.05, confirming statistical significance (Fig. 3E). Furthermore, a paired *t*-test of pART compared to *SymAori* segregation events resulted in $P < 0.05$ ($n = 20$; data not shown). Monitoring pART-eYFP in a strain expressing a plasmid-borne *parB-mCherry* fusion for labelling of *Cori* confirmed that, referred to segregation of *Cori*, segregation of pART and *SymAori* follow similar spatiotemporal patterns (Fig. S3C and Supporting Information Video S12). These consistencies indicate that the *repABC* cassette is sufficient to govern

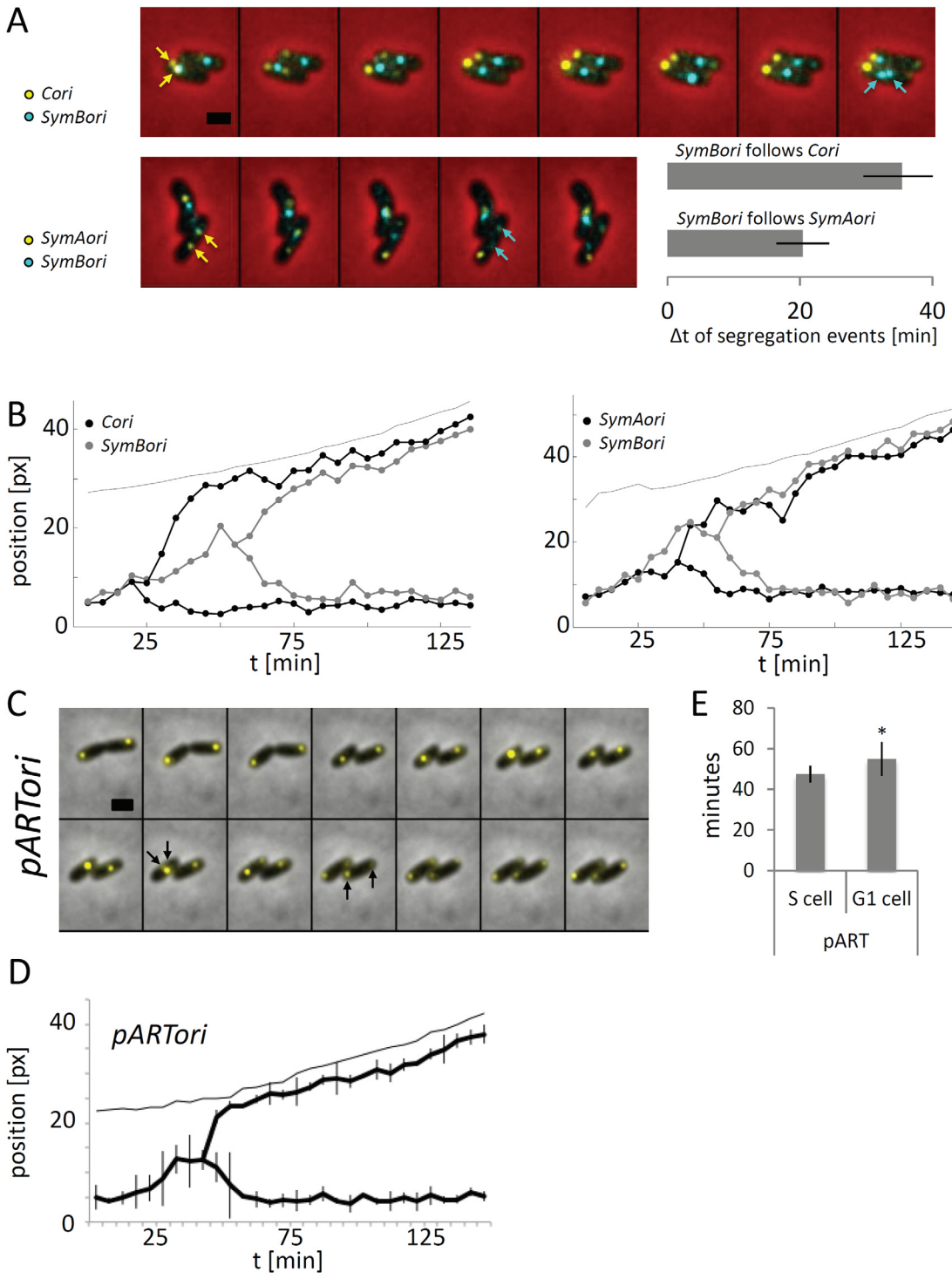


Fig. 3. Dual-colour origin tracking of the *S. meliloti* replicons.
A. Representative time-lapse fluorescence images of cells carrying *Cori*-eYFP and *SymBori*-eCFP, or *SymAori*-eYFP and *SymBori*-eCFP FROS-tagged origins. Durations between segregation events are shown (see arrows, each $n = 12$). Five minute time points, scale bar 1 μm.
B. *Cori*-eYFP and *SymBori*-eCFP, or *SymAori*-eYFP and *SymBori*-eCFP tracking. Upper grey thin lines display cell lengths. Dots represent relative positions of origins within the cell at every 5 min over a whole division cycle. Position [px] 0 = old pole.
C. Representative time-lapse fluorescence microscopy images of cells carrying *pARTori*-eYFP. Arrows show segregation events. Five minute time points, scale bar 2 μm.
D. *pARTori*-eYFP tracking. Illustration as in Fig. 2B.
E. Statistically significant (*$P < 0.05$) temporal order of segregation events after cell division of a strain cured of pSymA, carrying the 11.3 kb replicon pART: S cell 47.5 ± 5.4 min, G1 cell 55 ± 6.7 min. Error bars represent temporal variation, $n = 24$.

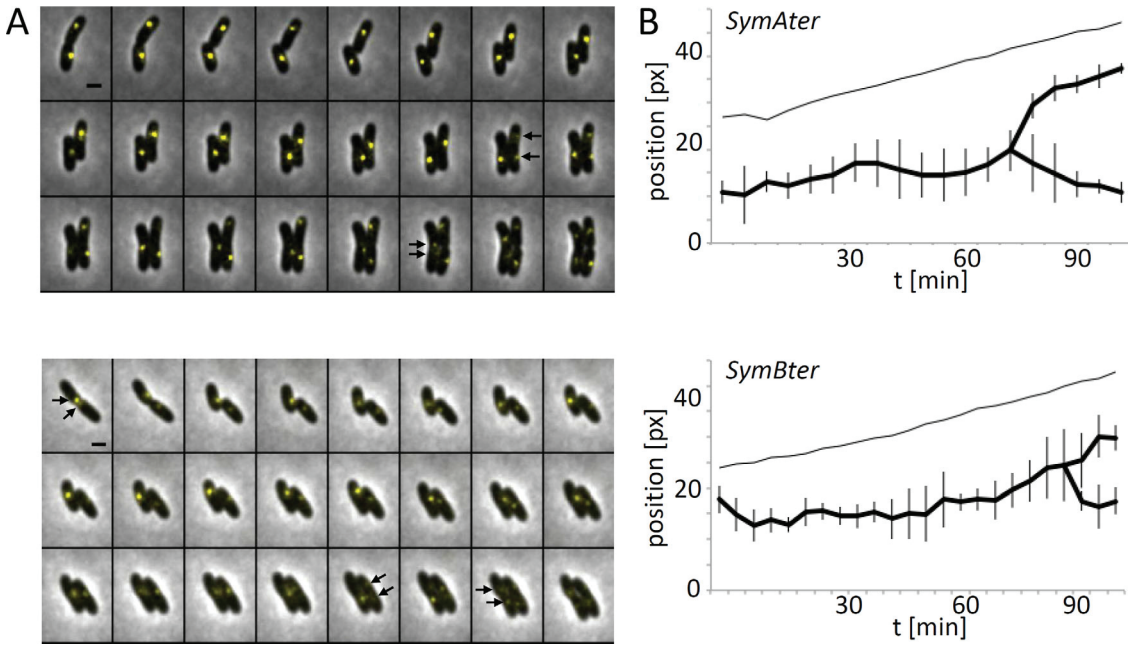


Fig. 4. Spatiotemporal dynamics of megaplasmid termini.
A. Representative time-lapse fluorescence microscopy images of cells carrying *SymAter*-eYFP or *SymBter*-eYFP FROS labels. Shown are overlays of fluorescence and phase contrast channels of dividing cells until origin segregation events were observed (see arrows). Five minute time points, scale bar 1 μm.
B. Tracking of termini regions. Upper thin lines display cell lengths. Thick lines represent medians of relative origin positions ($n = 4$) within the cell at every 5 min over a whole division cycle. Position [px] 0 = old pole.

DNA replication only once per cell cycle and for segregation of the origin following native timing.

C. crescentus *and* S. meliloti *Hda likely fulfill similar functions in regulating DnaA activity*

DnaA has previously been confirmed to function as replication initiator of the *S. meliloti* chromosome and binding of this protein to *dnaA* boxes within the chromosomal origin region located between *hemE* and *SMc02793* has been reported. Furthermore, DnaA was found to form a complex with the *repA2* promoter region of pSymA which contains a predicted DnaA box (Sibley *et al.*, 2006). This led us to ask whether DnaA and Hda play a role in coordinating initiation of chromosome and secondary replicons in *S. meliloti*.

We identified an *S. meliloti* Hda homologue (SMc00617) sharing 39% identity with the *C. crescentus* HdaA protein. *S. meliloti* Hda also shows 27% identity to the *S. meliloti* DnaA protein, with the homologous region partly including the predicted ATPase domain III of DnaA (Supporting Information Fig. S4A).

In agreement with a role of Hda in control of DnaA activity in *S. meliloti*, induced overexpression of *hda* from plasmid pSRK-Hda affected cell morphology and DNA content. Cells became homogeneously elongated (Fig. 5A) and a time course of cell lengths measured every 3 h after induction showed continuous cell elongation during the first 9 h (Supporting Information Fig. S4B). Constitutive overproduction of *S. meliloti* DnaA was previously shown to lead to filamentous or branched cells (Sibley *et al.*, 2006). Similar heterogeneous cell morphologies were observed 9 h after induction of *dnaA* overexpression from plasmid pSRK-DnaA (Fig. 5A). DNA content was measured by FACS analysis 0, 3, 6 and 9 h after induction of the overexpression constructs. Whilst the DNA content continuously increased during DnaA overproduction, accumulation of DNA was arrested at about 3–6 h after induction of *hda* overexpression (Fig. 5B). In summary, characterisation of *S. meliloti* Hda suggests that, as in *C. crescentus*, this protein likely is involved in regulatory inhibition of DnaA.

Altering DnaA activity uncouples replication control of chromosome and megaplasmids

To dissect whether the effect of DnaA activity on DNA content was replicon-specific, we determined the DNA content specific to each replicon applying comparative genomic hybridisation (CGH) (Fig. 5C). Comparison of cells 6 h after induction of DnaA overproduction to an uninduced control showed accumulation of chromosomal DNA relative to pSymA and pSymB content. In contrast, 6 h after induction of *hda* overexpression the

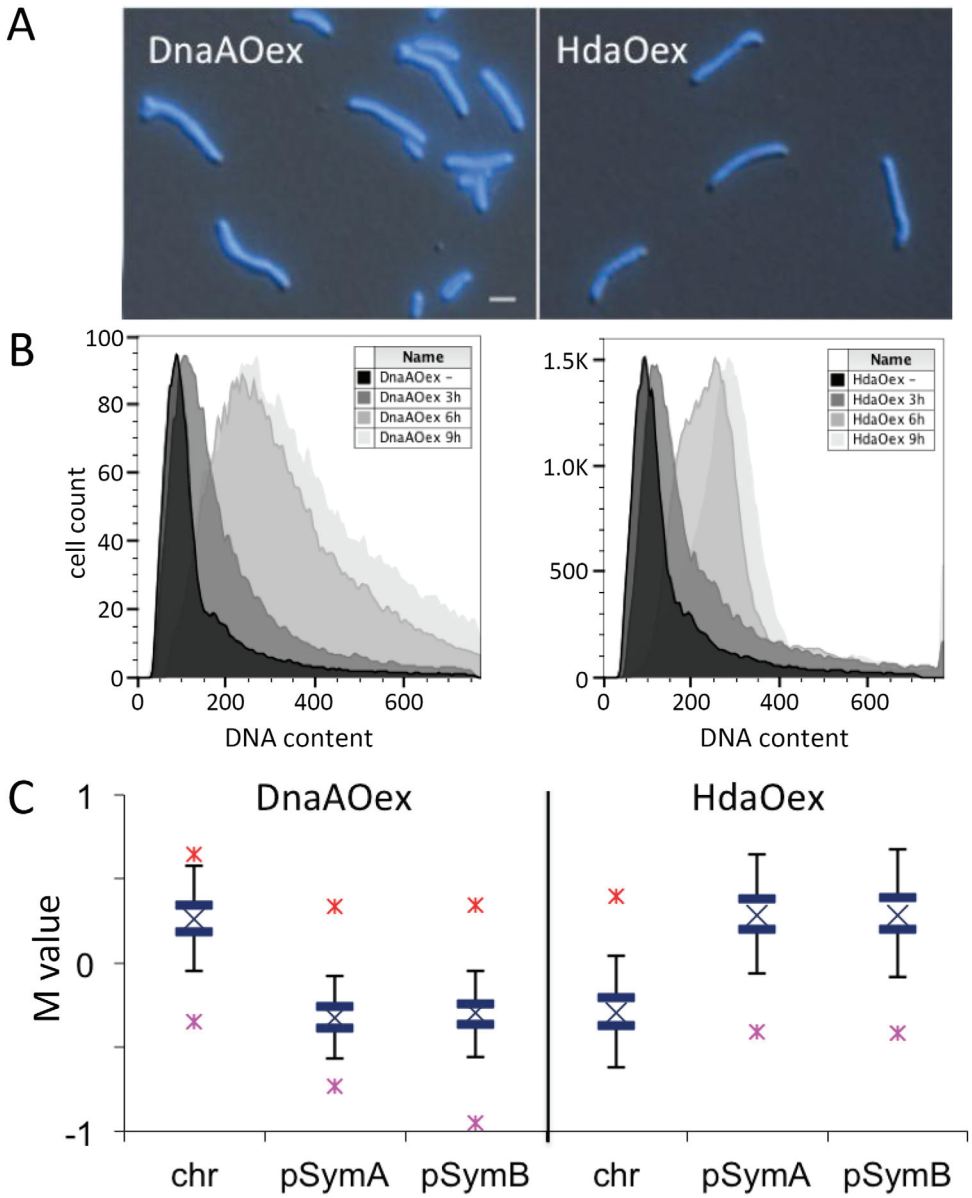


Fig. 5. Influence of DnaA activity on replicon content.
A. Morphological phenotype of cells carrying pSRK-DnaA (DnaAOex) or pSRK-Hda (HdaOex) 6 h after induction of *dnaA* or *hda* overexpression. Shown are overlays of DIC and DAPI stained DNA channels. Scale bar 2 μm.
B. Time course FACS analysis of DNA content of DnaAOex and HdaOex 3, 6 and 9 h after induction.
C. Comparative genomic hybridisation of DnaAOex and HdaOex strains induced for 6 h vs. uninduced controls. M values indicate relative DNA abundance of induced cultures, ordered by replicons. X represents median, thick blue lines represent upper and lower quartile, error bars indicate upper and lower whisker and asterisks indicate outlaws.

bias was shifted to the contrary, with less chromosomal DNA in relation to the other replicons.

FROS-labelled origins of chromosome and megaplasmids were counted in DnaA or Hda overproducing strains (Fig. 6). Upon DnaA overexpression, 41% of *Cori*-tagged cells showed more than two, and up to six fluorescent foci, whilst *SymAori*- or *SymBori*-tagged cells displayed one or two foci. Six hours after induction of *hda* overexpression, 92% of the *Cori*-tagged cells exhibited only a single fluorescent focus, whilst the number of one or two *SymAori*- and *SymBori*-mediated foci was little affected. Taken together, these results suggest that affecting DnaA activity, either positively or negatively, has a direct impact on the chromosome replication initiation only, leading to chromosomal overinitiation or G1 arrest respectively but does not affect replication of the megaplasmids.

Discussion

Organisms with divided genomes need to coordinate timing of once-per-cell-cycle replication and faithful segregation in order to avoid interference amongst multiple replicons. Whereas eukaryotic chromosome replication involves universal regulators, bacteria with multipartite genomes investigated so far rely on replicon-specific regulators (Kawakami and Katayama, 2010). Mechanisms controlling replication initiation of bacterial secondary chromosomes and megaplasmids, and

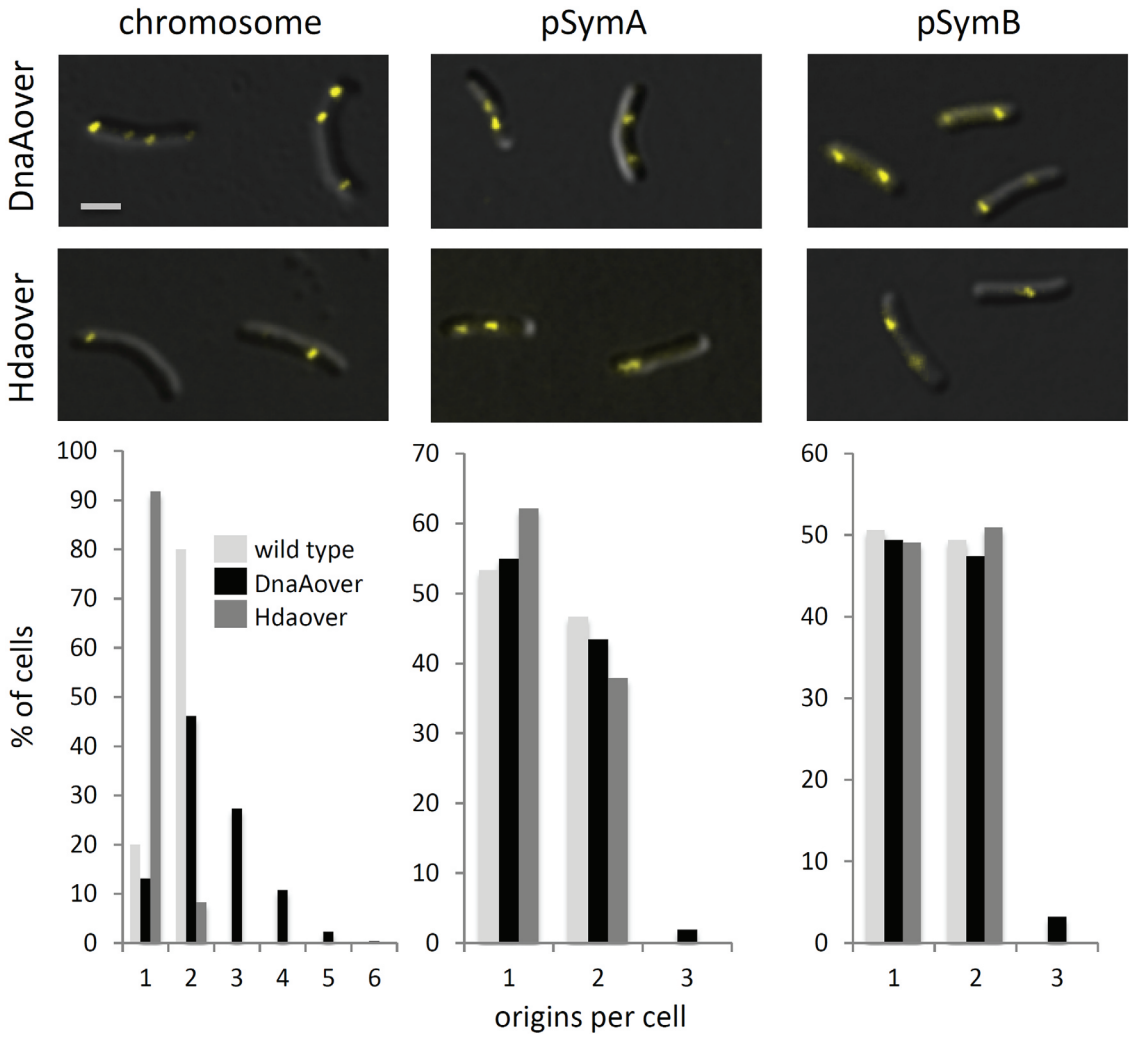


Fig. 6. Origin abundance after DnaA or Hda overproduction. *Cori*-eYFP, *SymAori*-eYFP and *SymBori*-eYFP foci were observed 6 h after pSRK-DnaA or pSRK-Hda mediated overexpression. Shown are overlays of fluorescent images of YFP and DIC channels. $n > 250$ cells for each construct. Scale bar 1 μm.

integration of these processes into the core cell cycle are just getting uncovered (Kahng and Shapiro, 2003; Cervantes-Rivera, 2011; Val *et al.*, 2014; Baek and Chattoraj, 2014).

In *S. meliloti*, we unraveled a temporally and spatially highly ordered choreography of segregation events of the tripartite genome. Whilst asymmetric cell division, initial chromosomal origin dynamics, initial replisome formation at the old pole, and G1 arrest of the smaller daughter cell are striking similarities shared by the cell cycles of the α-proteobacteria *C. crescentus* (Jensen *et al.*, 2001; Viollier *et al.*, 2004; Collier and Shapiro, 2009; Lin *et al.*, 2010) and *S. meliloti*, formation of multiple DnaN foci, indicative of multiple replisomes, are unique to *S. meliloti*. In contrast to *E. coli*, where the left-*ori*-right organisation of the chromosome results in two widely separated replisomes (Reyes-Lamothe *et al.*, 2008), in *C. crescentus* the distance between replisomes is shorter due to the *ori-ter* organisation of the chromosome (Viollier *et al.*, 2004) and a second replisome focus was only observed in 2% of analysed cells (Jensen *et al.*, 2001). In *S. meliloti*, formation of additional replisomes during the cell cycle is most likely associated to replication initiation of the two megaplasmids, and spatially and temporally separated from replication initiation of the chromosome. This is supported by (i) temporal correlation of occurrence of additional DnaN foci and duplicated megaplasmid origins, (ii) spatial separation of megaplasmid origins from the chromosomal origins and (iii) upregulation of the megaplasmid-borne *repABC* transcripts (De Nisco *et al.*, 2014) in later stages of the cell cycle.

In *C. crescentus*, the replisome is assembled at the old pole during the swarmer-to-stalked cell differentiation,

and gradually moves to midcell over the course of the replication cycle (Collier and Shapiro, 2009). Upon completion of S phase the replisome disassembles, but cells continue to grow in a linear fashion until division initiates, indicating a distinct G2-like phase as proposed in other bacteria (Laub *et al*., 2000). The first 40 min of *S. meliloti* replisome dynamics strongly resemble the moving DNA replication factory model described for *C. crescentus* (Jensen *et al*., 2001), where the two bidirectional replication forks of the single chromosome stay in close proximity to each other. After release and duplication of the chromosomal origin, translocation to the poles was performed at higher speed consistent with a recent description of changing origin velocities upon *C. crescentus* chromosome segregation (Shebelut *et al*., 2010). Trajectories of partitioning chromosome and megaplasmid origins were basically similar, except for the latter showing a longer polar release phase along with a longer distance of travelling away from the old pole during this phase. Furthermore, in contrast to the extreme polar positions of the chromosomal origins, the megaplasmid origins showed a subpolar localisation after completion of segregation, consistent with the previously reported subpolar localisation of FISH-labelled pSymA and pSymB origins (Kahng and Shapiro, 2003).

In *S. meliloti* we observed a strict order of individual replicon partitioning events, with the chromosome segregation starting first, followed by pSymA 20 min later, and by pSymB after further 15 min. Delayed replication initiation of secondary replicons was observed for chromosome 2 of *V. cholerae* (Rasmussen *et al*., 2007), and recently for chromosome 2 of the α-proteobacterial human and animal pathogen *Brucella abortus* (Deghelt *et al*., 2014). Although our data do not allow to determine replication initiation of all *S. meliloti* replicons, sequential segregation of the origins of replication of the chromosome and both megaplasmids is in agreement with a common strategy of temporal coordination of replication and segregation employed by bacteria carrying multipartite genomes.

In *V. cholerae* at the beginning of the cell cycle, both chromosomes are longitudinally oriented in the cell, with the origin of chromosome 1 located at the old and its terminus at the new pole whilst the origin of chromosome 2 resides at midcell and its terminus closer to the new pole (Egan *et al*., 2005; Fiebig *et al*., 2006; David *et al*., 2014). In *S. meliloti* at the start of the cell cycle, the origins of chromosome and megaplasmids localise to the old pole, whereas the terminus regions of pSymA and pSymB localise in the midcell area and close to the new pole respectively. In this respect, spatial organisation of *S. meliloti* megaplasmids and *V. cholerae* chromosomes seems to differ. In *V. cholerae*, initiation of the smaller chromosome 2 is delayed until two-third of the larger chromosome 1 has been replicated resulting in synchronised replication termination of both *V. cholerae* replicons (Rasmussen *et al*., 2007). In *S. meliloti*, our data show that segregation of the smaller pSymA begins after one-third of the cell cycle before partitioning of the larger pSymB, and that towards the end of the cell cycle only a single DnaN cluster remains. Hence, synchronisation of replication termination of all three replicons seems to be less probable. Segregation of the duplicated pSymB terminus regions started about 20 min after segregation of those of pSymA. The deduced time span between segregation of origin and terminus regions of pSymA was about 25% shorter than for pSymB, which correlates well to the 20% difference in size of these megaplasmids.

Little is known about the coordinated control of bacterial multireplicon replication. Recently, a binding site for the replication initiator protein RctB of *V. cholerae* chromosome 2 was identified on chromosome 1 (Baek and Chattoraj, 2014). A communication mechanism between both chromosomes was suggested based on remodeling of RctB by a chaperone-like activity of this binding site resulting in enhanced replication of chromosome 2. In *S. meliloti*, our data imply that segregation of the secondary replicons initiate after the origin of the earlier partitioning replicon (chromosome or megaplasmid) has reached the new pole. Therefore, it is tempting to speculate that origin tethering to the new pole may serve as a checkpoint to initiate the following replicon initiation or segregation events.

As *repABC* transcripts are upregulated later in the cell cycle (De Nisco *et al*., 2014) and have been shown to be involved in partitioning of the megaplasmids (Pinto *et al*., 2012) these proteins most likely drive the fast translocation of these replicons following origin duplication. The nucleoprotein complex ParABS is composed of the partitioning proteins ParA and ParB and the *parS* site located close to the origin of replication. In many bacteria, this complex is critical for chromosome segregation. How the components of the ParAB-*parS* or RepAB systems work together to generate translocation force is uncertain, yet a recent study describes a model where DNA-bound ParA-ATP dimers serve as transient tethers that use the elastic dynamics of the chromosome to relay the partition complex from one DNA region to another across a ParA-ATP dimer gradient (Lim *et al*., 2014). In this model, ParB intermittently binds to the DNA-bound ParA-ATP dimer, which probably adopts an ATP-hydrolysis competent state. A high ParB concentration (as present in the *parS*-ParB complex) is needed to stimulate ParA ATPase activity resulting in release from the DNA. This model may also apply to partitioning of megaplasmids involving their specific RepAB-*parS* complexes.

Specific to α-proteobacteria, CtrA serves as a master regulator of cell-cycle-dependent transcription and was shown to inhibit replication initiation by *Cori* sequestration in *C. crescentus* (Quon *et al*., 1998). *S. meliloti Cori* does not contain CtrA binding motifs (Brilli *et al*., 2010) and is not bound by CtrA, but depletion of CtrA led to an increase in cellular DNA content whilst maintaining the equal ratio of the three replicons in *S. meliloti* (Pini *et al*., 2015). Therefore, a mechanism upstream or independent of CtrA may be involved in communication between chromosome and secondary replicons for replication. Altering DnaA activity either positively or negatively by overexpression of *dnaA* or *hda* respectively resulted in an increase in replication initiation events or G1 arrest of the chromosome only, suggesting that DnaA or Hda are not essential components in this communication. Interestingly, the *repA2B2C2* cassette of pSymA alone was sufficient to confer the spatiotemporal properties of this megaplasmid to a small artificial plasmid in the *S. meliloti* cell cycle. Although the mechanisms of integration of the megaplasmids into the cell cycle remain unknown this observation certifies that all megaplasmid-encoded factors and/or cis-sequence elements required for these coordinated cell cycle processes are located within the *repABC* region.

Experimental procedures

Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 1 and oligonucleotides in Supporting Information Table S1. *S. meliloti* strains were grown in tryptone yeast (TY) full medium or in GMS minimal medium (Zevenhuizen and van Neerven, 1983) at 30°C, *E. coli* strains in LB medium at 37°C. When required, antibiotics were complemented to the following final concentrations (μg/l for solid media plates): streptomycin (Sm) 600; nalidixic acid (Nx) 10; kanamycin (Km) 200; gentamycin (Gm) 40 and tetracycline (Tc) 8 for *S. meliloti* strains. For *E. coli* strains Km 50 or Gm 10 were added. Final antibiotic concentrations were reduced to 50% in liquid cultures. pSRK-mediated overexpression (Khan *et al*., 2008) was induced by addition of IPTG to a final concentration of 0.5 mM. Twenty millimolar taurine was added to activate inducible FROS constructs 6 h prior to microscopic time-lapse observation.

Construction of S. meliloti *derivate strains*

Details of *S. meliloti* strain construction are provided in the supplementary material. The chromosomal *dnaN* coding region of *S. meliloti* wild-type 2011 and the pSymA-cured strain Sma818 (Oresnik *et al*., 2000) was replaced by a *dnaN-mCherry* translational fusion by double homologous recombination.

For live cell tracking of origins of replication, the two-color fluorescent repressor operator system (FROS) was applied (Lau *et al*., 2003). *tetO120* (120 *tet* operators) or *lacO120* (120 *lac* operators) cassettes were inserted into the chromosome (*dnaQ-seqB* intergenic region, 0° position), pSymA (*SMa2383-SMa2385* intergenic region, 0° position; *SMa1188-SMa1191* intergenic region, 180° position), or pSymB (*SMb20041-SMb20042* intergenic region, 1° position; *SMb21556* region, 233° position) by double homologous recombination.

The mini replicon pART comprises three functional modules. (i) The pSymA *repA2B2C2* region for propagation in *S. meliloti* and the *p15A* origin for maintenance in *E. coli*. Analog to MacLellan *et al*. (2006), for improved replicon stability the *repA2B2C2* region including only incy1 instead of incy1–5 sequences was used. (ii) A spectinomycin resistance cassette including the *aadA1* coding region. (iii) A 4.1 kb *tetO* array. pART was introduced to the pSymA-cured strain Sma818 (Oresnik *et al*., 2000) by electroporation.

For IPTG-induced overexpression, *dnaA* or *hda* coding regions were inserted into pSRK-Gm and introduced to either *S. meliloti* 2011 or FROS-tagged strains.

For operator array visualisation, taurine-induced P_{tauA} *lacI-eCFP tetR-eYFP* driven production of LacI-eCFP and TetR-eYFP proteins, specifically binding to *lacO* or *tetO* sites respectively was applied (Mostafavi *et al*., 2014; Lau *et al*., 2003). pK19m2 pTauA LacI-eCFP TetR-eYFP was integrated by single recombination into the genomic *tauA* promoter region of the *S. meliloti* strains.

Fluorescence microscopy

Bacteria were visually examined by differential interference contrast or phase contrast and epifluorescence or HILO fluorescence microscopy using a Nikon Eclipse Ti-E microscope equipped with 100× CFI Apo TIRF Oil objective (numerical aperture of 1.49) or 100× CFI Plan Apo Lambda DM Oil objective (NA of 1.45) with AHF HC filter sets F36-513 DAPI (excitation band pass 387/11 nm, beam splitter 409 nm, emission band pass 447/60 nm), F36-525 eGFP (excitation band pass 472/30 nm, beam splitter 495 nm, emission band pass 520/35 nm), F36-504 mCherry (excitation band pass 562/40 nm, beam splitter 593 nm, emission band pass 624/40 nm), F36-528 YFP (excitation band pass 500/24 nm, beam splitter 520 nm, emission band pass 542/27 nm) and F36-544 CFP (excitation band pass 438/24 nm, beam splitter 458 nm, emission band pass 483/32 nm). Fluorescence was excited by a multibeam argon laser for 457/488/514 nm, or a solid-state laser for 561 nm wavelengths. Images were acquired with an Andor iXon3 885 EMCCD camera. Exposure times ranged from 50 to 500 ms. Image acquisition and adjustment was done with Nikon NIS elements 4.0 software.

For time-lapse studies, overnight precultures grown in TY full medium with appropriate antibiotics were washed and resuspended in GMS minimal medium supplemented with 10% TY (GMSTY medium). FROS was induced by adding 20 mM taurine to the preculture 6 h prior observation. Logarithmic growing cells were placed on GMSTY 1% agarose pads and were allowed to adapt for 1 h before observation.

Table 1. Strains and plasmids.

Strain or plasmid	Relevant characteristics	Source
Strains		
S. meliloti 2011	Nx^r, Sm^r, Sm resistant derivative of *S. meliloti* SU47.	Casse (1979)
2011 *Cori*-eYFP *SymBori*-eCFP	2011 *Cori::tetO120* (0°) *SymBori::lacO120* (1°) *pTauA*::*lacI-eCFP tetR-eYFP*, Km^r	This work
2011 *Cori*-eYFP *SymBori*-eCFP	2011 *Cori::tetO120* (0°) *SymBori::lacO120* (1°) *pTauA*::*lacI-eCFP tetR-eYFP*, Km^r	This work
2011 *SymAori*-eYFP *SymBori*-eCFP DnaN-mCh	2011 *SymAori::tetO120* (0°) *SymBori::lacO120* (1°) *pTauA*::*lacI-eCFP tetR-eYFP*, DnaN replaced with *dnaN-mCherry*, Km^r, Gm^r	This work
2011 DnaN-mCh	2011 *dnaN* replaced with *dnaN-mCherry*, Gm^r	This work
2011 DnaAOex	2011 pSRKGm-DnaA	This work
2011 *Cori*-eYFP DnaAOex	2011 *Cori*-eYFP, pSRKGm-DnaA	This work
2011 *SymAori*-eYFP DnaAOex	2011 *SymAori*-eYFP, pSRKGm-DnaA	This work
2011 *SymBori*-eYFP DnaAOex	2011 *SymBori*-eYFP, pSRKGm-DnaA	This work
2011 HdaOex	2011 pSRKGm-Hda	This work
2011 *Cori*-eYFP HdaOex	2011 *Cori*-eYFP, pSRKGm-Hda	This work
2011 *SymAori*-eYFP HdaOex	2011 *SymAori*-eYFP, pSRKGm-Hda	This work
2011 *SymBori*-eYFP HdaOex	2011 *SymBori*-eYFP, pSRKGm-Hda	This work
Sma818	*S. meliloti* 2011 cured of megaplasmid pSymA	Oresnik *et al.* (2000)
Sma818 DnaN-mCh	Sma818 *dnaN* replaced with *dnaN-mCherry*, Gm^r	This work
S. meliloti pART	Sma818 harbouring pART mini-replicon	This work
S. meliloti pART ParB-mCherry	Sma818 harbouring pART mini-replicon, pSRK-ParB-mCherry	This work
2011 *pSymAter*-eYFP	2011 *SymAter::tetO120 pTauA*::*lacI-eCFP* tetR-eYFP, Km^r	This work
2011 *pSymBter-eYFP*	2011 *SymBter::tetO120 pTauA*::*lacI-eCFP* tetR-eYFP, Km^r	This work
Plasmids		
pK19mobsacB	Suicide plasmid in *S. meliloti*, *sacB*, *oriV*, Km^r	Schäfer *et al.* (1994)
pLAU43	pUC18 derivate, carrying 2 x 120 lac operator cassettes	Lau *et al.* (2003)
pLAU44	pUC18 derivate, carrying 2 x 120 tet operator cassettes	Lau *et al.* (2003)
pLAU53	pBAD24 derivate, carrying *lacI-eCFP*, *tetR-eYFP* translational fusions	Lau *et al.* (2003)
pMS252	pSVB28 carrying *aacC1*	Becker *et al.* (1995)
pMB393	pBBR1MCS carrying *aadA1*	Barnett *et al.* (2000)
pK19ms *Cori-tetO120*	pK19mobsacB carrying 120 tet operators and flanking regions for marker free insertion into *dnaQ-seqB* intergenic region	This work
pK19ms *SymAori-tetO120*	pK19mobsacB carrying 120 tet operators and flanking regions for marker free insertion into *SMa2383-2385* intergenic region	This work
pK19ms *SymBori-lacO120*	pK19mobsacB carrying 120 lac operators and flanking regions for marker free insertion into *SMb20041-20042* intergenic region	This work
pK19ms *SymBori-tetO120*	pK19mobsacB carrying 120 tet operators and flanking regions for marker free insertion into *SMb20041-20042* intergenic region	This work
pK19m2 pTauA LacI-eCFP TetR-eYFP	pk19mob2 taurine inducible *lacI-eCFP tetR-eYFP* expression construct, integrating into *pTauA* region	This work
pK19ms DnaN-mCherry	pK19mobsacB with *dnaN-mCherry* translational fusion, Gm^r, *dnaN* downstream region	This work
pSRKKm, pSRKGm	pBBR1MCS-2 derivative with a P_{lac} promoter, *lacIq*, *lacZa*$^+$, Km^r or Gm^r	Khan *et al.* (2008)
pSRKKm-ParB-mCherry	pRKKm carrying *S. meliloti parB* fused to *mCherry*	Elizaveta Krol
pSRKGm-DnaA	pSRKGm harbouring the *dnaA* coding region	Robledo *et al.* (2015)
pSRKGm-Hda	pSRKGm harbouring the *hda* coding region	This work
pUC18 mCh-*aacC1*	pUC18 carrying *mCherry* and *aacC1* coding regions	Greif *et al.* (2010)
pK18m2	Suicide vector, *mob*, *lacZ*, Km^r	Tauch *et al.* (1998)
pACYC177	*bla*, *p15A* origin, Km^r	Chang and Cohen (1978)
pJD45	pACYC177 derivate carrying pSymA *repA2B2C2*	This work
pART	artificial mini-replicon consisting of *repA2B2C2*, *incy*, *p15A*, *aadA1*, *tetO120*	This work

All experiments were performed at 30°C. Tracking of fluorescent tagged origins was performed using the Microbe Tracker Suite (Sliusarenko *et al.*, 2011). Time point zero was defined as the first image frame showing sibling cells physically separated after cell division.

Flow cytometry

Bacterial cultures were grown as indicated, and were prepared and analysed as previously described by Robledo *et al.* (2015).

Marker frequency analysis

S. meliloti was inoculated into 50 ml TY medium containing 600 μg/ml streptomycin at an initial cell density of OD_{600} of 0.1. After incubation at 30°C and 200 rpm, samples were taken at OD_{600} of 0.6 (exponential phase) or OD_{600} of 2.4 (overnight culture, stationary phase) and immediately frozen in liquid nitrogen. Preparation and acquisition of Illumina Miseq data were performed as previously described (Schäper *et al.*, 2015), obtaining the following reads: *S. meliloti* 2011 (stationary): adapter 1: 5.1×10^6, adapter 2: 3.1×10^6; *S. meliloti* 2011 (exponential): adapter 1: 4.8×10^6, adapter 2: 2.9×10^6

reads. Paired-end reads were mapped by the QuasR R package onto the *S. meliloti* replicons. Only unique hits were considered. Subsequently, chromosomal coverage was determined from the obtained mapped genomic DNA reads using the genomecov from the bedtools toolbox. Chromosomal coverage of exponential and stationary phase samples was normalised by the total coverage (sum of coverage) of each sample. To identify minimal variations in the copy number along the replicons, we used sliding window averaging. The size of the window comprises half of the chromosome. After averaging, the value at a certain chromosomal position reflects the average coverage of 25% of the chromosome left and right of the indicated position. This process averages out random noise. Finally, the ratio of exponential and stationary phase data sets was determined. To easily identify changes in the copy number, the ratio was further normalised such that the minimum (terminus region) was set to a copy number of one.

Comparative genomic hybridisation

Independent *S. meliloti* 2011 wild-type cultures grown in 50 ml TY rich medium were harvested in logarithmic phase (OD_{600} of 0.6) or stationary phase (overnight culture). *Sinorhizobium meliloti* 2011 carrying pSRKGm-DnaA or carrying pSRKGm-Hda were grown in 50 ml TY rich medium with appropriate antibiotics to OD_{600} of 0.6 and harvested 6 h after IPTG induction. Cells were disrupted and DNA was sheared by ultrasonication (Bioruptor, Diagenode). Total DNA was acquired by phenol/chloroform extraction, aminoallyl-dUTP was incorporated into the DNA by synthesis and labelled by Cy3- or Cy5-*succinimidyl ester*, followed by hybridisation of the labelled DNA to the Sm14kOLI microarray (ArrayExpress Accession No. A-MEXP-1760), image acquisition and data analysis as previously described (Serrania *et al*., 2008). The Sm14kOLI microarray carries 50 to 70mer oligonucleotides probes representing coding and intergenic regions of the *S. meliloti* 1021 genome. Normalisation and *t*-statistics were carried out using the EMMA 2.8.2 microarray data analysis software (Dondrup *et al*., 2009). Coding regions with P value ≤ 0.05 were included in the analyses, with the M value representing the $\log_2$ ratio of both channels. Microarray hybridisation data are available at ArrayExpress (accession number E-MTAB-3924).

Acknowledgements

We thank Bernadette Boomers, Andreas Kautz and Barbara Herte for technical assistance in whole genome sequencing and comparative genomics hybridisations, Elizaveta Krol for providing plasmid pSRKKm-ParB-mCherry, along with many other members of the Becker lab and SYNMIKRO for helpful discussions. This work was supported by the LOEWE Program of the State of Hesse (Germany).

References

Baek, J.H., and Chattoraj, D.K. (2014) Chromosome I controls chromosome II replication in *Vibrio cholerae*. *PLoS Genet* **10**: e1004184.

Barnett, M.J., Oke, V., and Long, S.R. (2000) New genetic tools for use in the Rhizobiaceae and other bacteria. *BioTechniques* **29**: 240–242, 244–245.

Becker, A., Schmidt, M., Jäger, W., and Pühler, A. (1995) New gentamicin-resistance and lacZ promoter-probe cassettes suitable for insertion mutagenesis and generation of transcriptional fusions. *Gene* **162**: 37–39.

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., and Biondi, E.G. (2000) The diversity and evolution of cell cycle regulation in alpha-proteobacteria: A comparative genomic analysis. *BMC Syst Biol* **4**: 52.

Cervantes-Rivera, R., Pedraza-Lopez, F., Perez-Segura, G., and Cevallos, M.A. (2011) The replication origin of a repABC plasmid. *BMC Microbiol* **11**: 158.

Collier, J., and Shapiro, L. (2009) Feedback control of DnaA-mediated replication initiation by replisome-associated HdaA protein in *Caulobacter*. *J Bacteriol* **191**: 5706–5716.

David, A., Demarre, G., Muresan, L., Paly, E., Barre, F.-X., and Possoz, C. (2014) The two Cis-acting sites, *parS1* and *oriC1*, contribute to the longitudinal organisation of *Vibrio cholerae* chromosome I. *PLoS Genet* **10**: e1004448.

De Nisco, N.J., Abo, R.P., Wu, C.M., Penterman, J., and Walker, G.C. (2014) Global analysis of cell cycle gene expression of the legume symbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci U S A* **111**: 3217–3224.

Deghelt, M., Mullier, C., Sternon, J.-F., Francis, N., Laloux, G., Dotreppe, D., *et al*. (2014) G1-arrested newborn cells are the predominant infectious form of the pathogen *Brucella abortus*. *Nat Commun* **5**: 4366.

del Solar, G., Giraldo, R., Ruiz-Echevarría, M.J., Espinosa, M., and Díaz-Orejas, R. (1998) Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**: 434–464.

Dondrup, M., Albaum, S.P., Griebel, T., Henckel, K., Jünemann, S., Kahlke, T., *et al*. (2009) EMMA 2 – A MAGE-compliant system for the collaborative analysis and integration of microarray data. *BMC Bioinform* **10**: 50.

Egan, E.S., Fogel, M.A., and Waldor, M.K. (2005) Divided genomes: Negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol Microbiol* **56**: 1129–1138.

Fiebig, A., Keren, K., and Theriot, J.A. (2006) Fine-scale time-lapse analysis of the biphasic, dynamic behaviour of the two *Vibrio cholerae* chromosomes. *Mol Microbiol* **60**: 1164–1178.

Fields, A.T., Navarrete, C.S., Zare, A.Z., Huang, Z., Mostafavi, M., Lewis, J.C., *et al*. (2012) The conserved polarity factor PodJ1 impacts multiple cell envelope-associated functions in *Sinorhizobium meliloti*. *Mol Microbiol* **84**: 892–920

Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., *et al*. (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**: 668–672.

González, V., Santamaría, R.I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., *et al*. (2006) The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci U S A* **103**: 3834–3839.

Greif, D., Pobigaylo, N., Frage, B., Becker, A., Regtmeier, J., and Anselmetti, D. (2010) Space- and time-resolved protein dynamics in single bacterial cells observed on a chip. *J Biotechnol* **149**: 280–288.

Hallez, R., Bellefontaine, A.-F., Letesson, J.-J., and De Bolle, X. (2004) Morphological and functional asymmetry in alpha-proteobacteria. *Trends Microbiol* **12**: 361–365.

Jensen, R.B., Wang, S.C., and Shapiro, L. (2001) A moving DNA replication factory in *Caulobacter crescentus*. *EMBO J* **20**: 4952–4963.

Jonas, K., Chen, Y.E., and Laub, M.T. (2011) Modularity of the bacterial cell cycle enables independent spatial and temporal control of DNA replication. *Curr Biol* **21**: 1092–1101.

Jonas, K., Liu, J., Chien, P., and Laub, M.T. (2013) Proteotoxic stress induces a cell-cycle arrest by stimulating Lon to degrade the replication initiator DnaA. *Cell* **154**: 623–636.

Kaguni, J.M. (2006) DnaA: Controlling the initiation of bacterial DNA replication and more. *Annu Rev Microbiol* **60**: 351–375.

Kahng, L.S., and Shapiro, L. (2003) Polar localization of replicon origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. *J Bacteriol* **185**: 3384–3391.

Katayama, T., Kubota, T., Takata, M., Akimitsu, N., and Sekimizu, K. (1996) Disruption of the *hslU* gene, which encodes an ATPase subunit of the eukaryotic 26S proteasome homolog in *Escherichia coli*, suppresses the temperature-sensitive *dnaA46* mutation. *Biochem Biophys Res Commun* **229**: 219–224.

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010) Regulation of the replication cycle: Conserved and diverse regulatory systems for DnaA and *oriC*. *Nat Rev Microbiol* **8**: 163–170.

Kato, J., and Katayama, T. (2001) Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*. *EMBO J* **20**: 4253–4262.

Kawakami, H., and Katayama, T. (2010) DnaA, ORC, and Cdc6: Similarity beyond the domains of life and diversity. *Biochem Cell Biol* **88**: 49–62.

Khan, S.R., Gaines, J., Roop, R.M., and Farrand, S.K. (2008) Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl Environ Microbiol* **74**: 5053–5062.

Kirkup, B.C., Chang, L., Chang, S., Gevers, D., and Polz, M.F. (2010) Vibrio chromosomes share common history. *BMC Microbiol* **10**: 137.

Kitagawa, R., Ozaki, T., Moriya, S., and Ogawa, T. (1998) Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for *Escherichia coli* DnaA protein. *Genes Dev* **12**: 3032–3043.

Lam, H., Matroule, J.-Y., and Jacobs-Wagner, C. (2003) The asymmetric spatial distribution of bacterial signal transduction proteins coordinates cell cycle events. *Dev Cell* **5**: 149–159.

Lau, I.F., Filipe, S.R., Søballe, B., Økstad, O.-A., Barre, F.-X., and Sherratt, D.J. (2003) Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol* **49**: 731–743.

Laub, M.T., McAdams, H.H., Feldblyum, T., Fraser, C.M., and Shapiro, L. (2000) Global analysis of the genetic network controlling a bacterial cell cycle. *Science* **290**: 2144–2148.

Lim, H.C., Surovtsev, I.V., Beltran, B.G., Huang, F., Bewersdorf, J., and Jacobs-Wagner, C. (2014) Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. *Elife* **3**: e02758.

Lin, Y., Crosson, S., and Scherer, N.F. (2010) Single-gene tuning of *Caulobacter* cell cycle period and noise, swarming motility, and surface adhesion. *Mol Syst Biol* **6**: 445.

Lu, M., Campbell, J.L., Boye, E., and Kleckner, N. (1994) SeqA: A negative modulator of replication initiation in *E. coli*. *Cell* **77**: 413–426.

MacLellan, S.R., Zaheer, R., Sartor, A.L., MacLean, A.M., and Finan, T.M. (2006) Identification of a megaplasmid centromere reveals genetic structural diversity within the *repABC* family of basic replicons. *Mol Microbiol* **59**: 1559–1575.

Mostafavi, M., Lewis, J.C., Saini, T., Bustamante, J.A., Gao, I.T., Tran, T.T., *et al.* (2014) Analysis of a taurine-dependent promoter in *Sinorhizobium meliloti* that offers tight modulation of gene expression. *BMC Microbiol* **14**: 295.

Mott, M.L., and Berger, J.M. (2007) DNA replication initiation: Mechanisms and regulation in bacteria. *Nat Rev Microbiol* **5**: 343–354.

Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J.A., *et al.* (2001) Complete genome sequence of *Caulobacter crescentus*. *Proc Natl Acad Sci U S A* **98**: 4136–4141.

Okada, K., Iida, T., Kita-Tsukamoto, K., and Honda, T. (2005) Vibrios commonly possess two chromosomes. *J Bacteriol* **187**: 752–757.

Oresnik, I.J., Liu, S.L., Yost, C.K., and Hynes, M.F. (2000) Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. *J Bacteriol* **182**: 3582–3586.

Panis, G., Murray, S.R., and Viollier, P.H. (2015) Versatility of global transcriptional regulators in alpha-Proteobacteria: From essential cell cycle control to ancillary functions. *FEMS Microbiol Rev* **39**: 120–133.

Pini, F., De Nisco, N.J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., *et al.* (2015) Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*. *PLoS Genet* **11**: e1005232.

Pini, F., Frage, B., Ferri, L., De Nisco, N.J., Mohapatra, S.S., Taddei, L., *et al.* (2013) The DivJ, CbrA and PleC system controls DivK phosphorylation and symbiosis in *Sinorhizobium meliloti*. *Mol Microbiol* **90**: 54–71.

Pinto, U.M., Pappas, K.M. and Winans, S.C. (2012) The ABCs of plasmid replication and segregation. *Nat Rev Microbiol* **10**: 755–765.

Quon, K.C., Yang, B., Domian, I.J., Shapiro, L., and Marczynski, G. (1998) Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosomal origin. *Proc Natl Acad Sci U S A* **95**: 120–125.

Rasmussen, T., Jensen, R.B., and Skovgaard, O. (2007) The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *EMBO J* **26**: 3124–3131.

Reyes-Lamothe, R., Possoz, C., Danilova, O., and Sherratt, D.J. (2008) Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell* **133**: 90–102.

Robledo, M., Frage, B., Wright, P.R., and Becker, A. (2015) A stress-induced small RNA modulates alpha-rhizobial cell cycle progression. *PLoS Genet* **11**: e1005153.

Schäper, S., Krol, E., Skotnicka, D., Kaever, V., Hilker, R., Søgaard-Andersen, L., and Becker, A. (2015) Cyclic di-GMP regulates multiple cellular functions in the symbiotic alphaproteobacterium *Sinorhizobium meliloti*. *J Bacteriol* **198**: 521–535.

Serrania, J., Vorhölter, F.-J., Niehaus, K., Pühler, A., and Becker, A. (2008) Identification of *Xanthomonas campestris* pv. campestris galactose utilization genes from transcriptome data. *J Biotechnol* **135**: 309–317.

Shebelut, C.W., Guberman, J.M., van Teeffelen, S., Yakhnina, A.A., and Gitai, Z. (2010) *Caulobacter* chromosome segregation is an ordered multistep process. *Proc Natl Acad Sci U S A* **107**: 14194–14198.

Sibley, C.D., MacLellan, S.R., and Finan, T. (2006) The *Sinorhizobium meliloti* chromosomal origin of replication. *Microbiology* **152**: 443–455.

Sliusarenko, O., Heinritz, J., Emonet, T., and Jacobs-Wagner, C. (2011) High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol Microbiol* **80**: 612–627.

Soufo, C.D., Soufo, H.J.D., Noirot-Gros, M.-F., Steindorf, A., Noirot, P., and Graumann, P.L. (2008) Cell-cycle-dependent spatial sequestration of the DnaA replication initiator protein in *Bacillus subtilis*. *Dev Cell* **15**: 935–941.

Su'etsugu, M., Nakamura, K., Keyamura, K., Kudo, Y., and Katayama, T. (2008) Hda monomerization by ADP binding promotes replicase clamp-mediated DnaA-ATP hydrolysis. *J Biol Chem* **283**: 36118–36131.

Val, M.-E., Soler-Bistué, A., Bland, M.J., and Mazel, D. (2014) Management of multipartite genomes: The *Vibrio cholerae* model. *Curr Opin Microbiol* **22**: 120–126.

Veening, J.-W., Murray, H., and Errington, J. (2009) A mechanism for cell cycle regulation of sporulation initiation in *Bacillus subtilis*. *Genes Dev* **23**: 1959–1970.

Viollier, P.H., Thanbichler, M., McGrath, P.T., West, L., Meewan, M., McAdams, H.H., and Shapiro, L. (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc Natl Acad Sci U S A* **101**: 9257–9262.

Wang, Q.P., and Kaguni, J.M. (1987) Transcriptional repression of the *dnaA* gene of *Escherichia coli* by dnaA protein. *Mol Gen Genet* **207**: 518–525.

Zevenhuizen, L.P.T.M., and van Neerven, A.R.W. (1983) (1-2)-β-d-Glucan and acidic oligosaccharides produced by *Rhizobium meliloti*. *Carbohydr Res.* **118**:127–134.

Zweiger, G., and Shapiro, L. (1994) Expression of *Caulobacter dnaA* as a function of the cell cycle. *J Bacteriol* **176**: 401–408.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

6.3 Cloning-free genome engineering in *Sinorhizobium meliloti* advances applications of Cre/*loxP* site-specific recombination

Eigener Beitrag zur Studie:

I. Planung der Experimente und Rekombinationsstrategien

II. Herstellung aller Konstrukte und Stämme

III. Durchführung aller Experimente mit Ausnahme des *lox spacer* Rekombinations-Tests

IV. Mitverfasser des Manuskripts

Journal of Biotechnology 233 (2016) 160–170

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Cloning-free genome engineering in *Sinorhizobium meliloti* advances applications of Cre/*loxP* site-specific recombination

Johannes Döhlemann, Meike Brennecke, Anke Becker*

LOEWE Center for Synthetic Microbiology and Faculty of Biology, Philipps-Universität Marburg, Marburg, Germany

ARTICLE INFO

Article history:
Received 19 May 2016
Received in revised form 26 June 2016
Accepted 30 June 2016
Available online 5 July 2016

Keywords:
Sinorhizobium meliloti
Electroporation
Cre/*loxP* recombination
Genome editing

ABSTRACT

The soil-dwelling α-proteobacterium *Sinorhizobium meliloti* serves as model for studies of symbiotic nitrogen fixation, a highly important process in sustainable agriculture. Here, we report advancements of the genetic toolbox accelerating genome editing in *S. meliloti*. The *hsdMSR* operon encodes a type-I restriction-modification (R-M) system. Transformation of *S. meliloti* is counteracted by the restriction endonuclease HsdR degrading DNA which lacks the appropriate methylation pattern. We provide a stable *S. meliloti hsdR* deletion mutant showing enhanced transformation with *Escherichia coli*-derived plasmid DNA and demonstrate that using an *E. coli* plasmid donor, expressing *S. meliloti* methyl transferase genes, is an alternative strategy of increasing the transformation efficiency of *S. meliloti*. Furthermore, we devise a novel cloning-free genome editing (CFGE) method for *S. meliloti*, *Agrobacterium tumefaciens* and *Xanthomonas campestris*, and demonstrate the applicability of this method for intricate applications of the Cre/*lox* recombination system in *S. meliloti*. An enhanced Cre/*lox* system, allowing for serial deletions of large genomic regions, was established. An assay of *lox* spacer mutants identified a set of *lox* sites mediating specific recombination. The availability of several non-promiscuous Cre recognition sites enables simultaneous specific Cre/*lox* recombination events. CFGE combined with Cre/*lox* recombination is put forward as powerful approach for targeted genome editing, involving serial steps of manipulation to expedite the genetic accessibility of *S. meliloti* as chassis.

1. Introduction

Although nitrogen gas is the main component of the atmosphere, it cannot be directly assimilated by plants. Agricultural practices therefore require an effective management of nitrogen availability. Modern agriculture faces challenges, such as excessive use, leaching and volatilization of mineral nitrogen fertilizer that is produced in an energy intensive process. Because of this environmental burden, leguminous plant crops fixing nitrogen by means of endosymbiotic rhizobia increasingly gain importance in agricultural practices.

Recent technical advances in life sciences enable extensive genomic manipulations and promote synthetic biology in designing and constructing biological systems towards a better understanding and diverse applications. These technical advances as well as an increasing understanding of the molecular principles of prokaryotic nitrogen fixation and the rhizobia-legume symbiosis have inspired ambitious long-term projects to engineer the nitrogen fixation pathway and to introduce this pathway or nitrogen fixing bacteria into cereal crops (Mus et al., 2016). These approaches are complex engineering problems, but have the potential to provide biotechnological solutions to the nitrogen problem.

Despite the increasing importance of nitrogen-fixing rhizobia, the genetic toolbox for engineering of α-proteobacterial rhizobia is not well developed regarding the demands of recent synthetic biology approaches. *Sinorhizobium meliloti*, the microsymbiont of plants of the genera *Medicago*, *Melilotus* and *Trigonella*, is a well-studied model for symbiotic nitrogen fixation (Gibson et al., 2008). The *S. meliloti* genome is commonly composed of a chromosome (3.65 Mb) and the megaplasmids pSymA (1.35 Mb) and pSymB (1.68 Mb) (Galibert et al., 2001). Natural populations of this species show extensive variations at genetic level, with the highest variation observed on the megaplasmids (Galardini et al., 2011; Giuntini et al., 2005). Also the genome architecture was found to be highly dynamic. Strains with various genome configurations and even fusions of the chromosome and both megaplasmids were identified (Guo et al., 2003). Construction of strains characterized by large deletions on the megaplasmids, or even cured for one or both of these replicons (diCenzo et al., 2014; Oresnik et al., 2000), further underscore the robustness of *S. meliloti* against a wide range

* Corresponding author.
E-mail address: anke.becker@synmikro.uni-marburg.de (A. Becker).

http://dx.doi.org/10.1016/j.jbiotec.2016.06.033
0168-1656/

of genomic modifications. This robustness and efficient methods for genome engineering are important qualities of a chassis in synthetic biology.

A variety of powerful techniques are available for genome engineering in *S. meliloti*, including the lambda integrase, Flp/FRT and Cre/*loxP* recombination systems (Harrison et al., 2011; House et al., 2004; Milunovic et al., 2014). The bacteriophage tyrosine recombinase Cre catalyzes reciprocal non-homologous recombination at a specific DNA recognition site (*loxP*) (Abremski and Hoess, 1984; Hoess et al., 1982; Sternberg and Hamilton, 1981). This 34 bp element consists of 13 bp palindromic sequences (left and right arm) flanking an 8 bp spacer region. Both left/right arms and spacer region have been mutated, making the Cre reaction largely unidirectional and specific (Araki et al., 1997; Lee and Saito, 1998; Langer et al., 2002; Missirlis et al., 2006). Proper functioning of Cre-mediated recombination is independent of host cofactors or accessory proteins. Genome engineering applications based on Cre/*loxP* were successfully employed in various bacterial model organisms (Lambert et al., 2007; Suzuki et al., 2005; Yu et al., 2002), including *S. meliloti* (Harrison et al., 2011) and *A. tumefaciens* (Liu et al., 2016). Since advanced genome engineering approaches rely on serial genomic integration of functional elements, common cloning procedures pose a significant hindrance in Cre/*loxP* applications in terms of speed and multiple usage of genetic elements, such as plasmid vectors. To fully exploit the potential of this recombination system for application in comprehensive genome engineering and to accelerate simple manipulations (e.g. gene knockout or tagging), a fast-track procedure for genome editing would be of great benefit.

For transformation of *S. meliloti* the method of choice is *Escherichia coli*-mediated conjugal transfer of plasmid DNA. But from the viewpoint of genome engineering conjugation imposes several restrictions. Common protocols are time-consuming and based on only a few plasmid families. Electroporation can overcome most drawbacks associated with conjugation. However, electroporation of plasmid DNA into wild type bacteria is mostly inefficient, which is frequently caused by restriction-modification (R-M) systems. Whereas most *Sinorhizobium* species are supposed to lack an R-M system, such a system encoded by the chromosomal *hsdMSR* operon was identified in the model organism *S. meliloti* Rm1021 (Capela et al., 2001; Ferri et al., 2010). In type-I R-M systems the role of HsdR is to cleave DNA lacking the appropriate methylation pattern. HsdS and HsdM form a complex in which HsdS is required for specific DNA recognition, and HsdM binds and carries out the methylation of adenosine residues (Ferri et al., 2010; Murray, 2000). Inactivation of the putative type-I restriction enzyme HsdR resulted in elevated transformation rates of *S. meliloti* Rm1021 (Ferri et al., 2010).

In this study, we advance the genetic toolbox for *S. meliloti*. We provide a stable *S. meliloti hsdR* deletion mutant showing enhanced transformation with *E. coli*-derived plasmid DNA, and demonstrate that using an *E. coli* plasmid donor expressing *S. meliloti hsdMS-ccrM* is an alternative strategy of increasing the transformation efficiency of *S. meliloti*. Furthermore, we devise a cloning-free genome editing (CFGE) procedure, show that this method is also applicable in the related α-proteobacterium *A. tumefaciens* and the γ-proteobacterium *X. campestris*, and pave the way for intricate and serial applications of the Cre/*lox* recombination system in *S. meliloti*.

2. Material and methods

2.1. Bacterial strains and growth conditions

Bacterial strains used in this study are derivatives of *E. coli* K12, *S. meliloti* Rm1021, *A. tumefaciens* C58 and *X. campestris* B100 (*X. campestris*) (Table 1). *E. coli* strains were grown at 37 °C in Luria-Bertani (LB) medium (Sambrook and Russel, 2001). *S. meliloti* strains were cultivated at 30 °C in tryptone yeast extract (TY) medium (Beringer, 1974). *A. tumefaciens* and *X. campestris* strains were grown at 30 °C in LB medium. If necessary, media were supplemented with the following antibiotics: Ampicillin (Amp; 100 μg/ml for *E. coli*), gentamicin (Gm; 10 μg/ml for *E. coli*, 20 μg/ml for *S. meliloti* and 30 μg/ml for *X. campestris* and *A. tumefaciens*), kanamycin (Km; 50 μg/ml for *E. coli* and 100 μg/ml for *S. meliloti*), streptomycin (Str; 600 μg/ml for *S. meliloti* and 800 μg/ml for *X. campestris* and *A. tumefaciens*), or tetracycline (Tc; 5–10 μg/ml for *S. meliloti*). Calcofluor fluorescence assays were performed on TY (*S. meliloti*) or LB (*X. campestris*, *A. tumefaciens*) agar plates supplemented with appropriate antibiotics and 200 μg/ml calcofluor (Fluorescent Brightener 28, Sigma). Cell suspensions were adjusted to an initial cell density at 600 nm (OD_{600}) of 0.1, dropped on agar plates and incubated over 48 h at 30 °C.

2.2. Plasmid and strain construction

Strains and constructs used in this study are listed in Table 1. Oligonucleotides are listed in Table S2 and a detailed description of constructs is given in Table S1.

2.2.1. Construction of strains and plasmids for transformation studies

For construction of *S. meliloti* strain SmΔhsdR, DNA fragments flanking the coding sequence of *hsdR* (SMc02292) were PCR-amplified and inserted into suicide vector pK18mobsacB, resulting in pJD11. The construct was introduced into *S. meliloti* and the backbone was subsequently deleted via sucrose selection (Schäfer et al., 1994). In order to determine integration efficiency of a typical construct for genome engineering applications, ~500 bp homologous sequences for integration into different loci on pSymB were cloned into a pK18mob2 derivative which was designed for Cre/*loxP* procedures, giving rise to pJD23-25 ("Insert 1–3"). In order to introduce DNA methyltransferase genes into *E. coli*, the moderate-copy number vector pSRK and the low-copy number vector pPHU231 were used. The coding sequence (CDS) of CcrM (SMc00021) was integrated into pSRKKm, giving rise to pSRKKm:*ccrM*. Sequences of *hsdM* (SMc02296) and *hsdMS* (SMc02295-02296) including the native RBS were inserted into pSRKGm, resulting in constructs pJD29 and pJD30, respectively. For construction of a pPHU231-based helper plasmid *hsdMS* and *ccrM* were cloned under the control of the PBAD promoter (Guzman et al., 1995), giving rise to pJD109. The construct was introduced into *E. coli* DH5α, resulting in strain DH5α109.

2.2.2. Construction of strains and plasmids for Cre/lox applications

In order to construct an *S. meliloti cre* expression strain a DNA fragment comprising a sequence for homologous recombination with the *tauX* locus and the coding sequences of Cre recombinase and tetracycline resistance cassette was released from pJG468 and inserted into pK18mobsacB. Cloning of a second region for homologous recombination with the *tauX* locus resulted in pJD19 and enabled deletion of the plasmid backbone via sucrose selection (Schäfer et al., 1994). The resulting marker-free strain SmCre was also transformed with pJD11 for deletion of *hsdR*, giving rise to strain SmCreΔhsdR. Basically, all constructs produced to assay and validate Cre/*lox* applications mediate deletion of the *exo* gene cluster on pSymB by use of different *lox* site derivatives. Plasmids pJD35, pJD78, pJD79, pJD80 were used for determination of Cre-mediated recombination efficiency resulting from differently mutated *lox* sites. Both pJD35 and pJD78 are based on pK18mobsacB carrying a native *loxP* site and *loxR*, respectively. *lox* sites are flanked by

Table 1
Bacterial strains and plasmids used in this study.

Strains/Plasmids	Description	References
S. meliloti		
Rm1021	Wild type strain	Galibert et al. (2001)
B033	Rm1021 carrying excision module for Cre/*loxP* studies	Harrison et al. (2011)
SmΔhsdR	Rm1021 *hsdR* deletion strain	This work
SmCre	Rm1021 *cre* expression strain, *tauX*:*cre-tetRA*	This work
SmCreΔhsdR	Rm1021 *cre* expression strain with *hsdR* deletion, *tauX*: *cre-tetRA*	This work
Rm1021 exoB	CFGE-mediated inactivation of *exoB*	This work
JDSm80	SmCre prepared for *loxLR*/*loxP*-mediated deletion of the *exo* gene cluster	This work
JDSm84	SmCre prepared for *loxL*/*loxR*-mediated deletion of the *exo* gene cluster	This work
JDSm85	SmCreΔhsdR prepared for *loxL*/*loxR*-mediated deletion of the *exo* gene cluster	This work
JDSm93	JDSm85Δ*exo* gene cluster	This work
JDSm128	SmCreΔhsdR derivative used for mutant *lox* site assay (*lox1*)	This work
JDSm129	SmCreΔhsdR derivative used for mutant *lox* site assay (*lox2*)	This work
JDSm130	SmCreΔhsdR derivative used for mutant *lox* site assay (*lox3*)	This work
JDSm131	SmCreΔhsdR derivative used for mutant *lox* site assay (*lox4*)	This work
JDSm132	SmCreΔhsdR derivative used for mutant *lox* site assay (*lox5*)	This work
JDSm133	SmCreΔhsdR derivative used for mutant *lox* site assay (*loxP*)	This work
A. tumefaciens		
C58	Wild type strain	Goodner et al. (2001)
C58 exoB	CFGE-mediated inactivation of *exoB*	This work
X. campestris		
B100	Derivative of strain DSM1526	Vorhölter et al. (2008)
B100 exoB	CFGE-mediated inactivation of *exoB*	This work
E. coli		
DH5α	Used for cloning procedures and plasmid extraction	Hanahan (1983)
MG1655	Used for plasmid extraction	Blattner et al. (1997)
MC4100	Used for plasmid extraction	Peters et al. (2003)
S17-1	Used for conjugation with *S. meliloti*	Simon et al. (1983)
DH5α109	DH5α carrying helper plasmid pJD109	This work
Plasmids		
pBad18	Expression vector containing promoter PBAD (Amp^r)	Guzman et al. (1995)
pK18mob2	Suicide vector (Km^r)	Schäfer et al. (1994)
pK18mobsacB	Suicide vector carrying *sacB* for sucrose selection (Km^r)	Schäfer et al. (1994)
pSRK-Gm/Km	Broad-host-range expression vector (Gm^r, Km^r)	Khan et al. (2008)
pPHU231	Low-copy vector (Tc^r)	Hübner et al. (1991)
pJG468	P_{tau}-*cre* integration vector (Tc^r)	Harrison et al. (2011)
pJG181	Excision module vector (Gm^r)	Harrison et al. (2011)
pJD11	*hsdR* deletion construct (Km^r)	This work
pJD19	*cre-tetRA* integration vector (Km^r, Tc^r)	This work
pJD23	pK18mob2 derivative carrying ~550 bp *exoP* fragment (Km^r)	This work
pJD24	pK18mob2 derivative carrying ~550 bp *exoI* fragment (Km^r)	This work
pJD25	pK18mob2 derivative carrying ~550 bp *exsA* fragment (Km^r)	This work
pSRKKm:*ccrM*	pSRKKm carrying *ccrM* (Km^r)	B. Frage (unpublished)
pJD29	pSRKGm carrying *hsdM* (Gm^r)	This work
pJD30	pSRKGm carrying *hsdMS* (Gm^r)	This work
pJD109	pPHU231 carrying *hsdMS-ccrM* (Tc^r)	This work
pJD35	pK18mobsacB carrying *loxP* (Km^r)	This work
pJD78	pK18mobsacB carrying *loxR* (Km^r)	This work
pJD79	pK18mobsacB carrying *loxL* (Km^r)	This work
pJD80	pK18mobsacB carrying *loxLR* (Km^r)	This work
pCreDelR	pK18mob2 carrying *aacC1* and *loxR* (Km^r/Gm^r)	This work
pCreDelL	pK18mobsacB carrying *loxL* (Km^r)	This work
pJD116	pCreDelL carrying a *lox* array providing *loxL* and *lox1L-lox5L* (Km^r)	This work
pJD121	pCreDelR carrying *lox1R* (Km^r/Gm^r)	This work
pJD122	pCreDelR carrying *lox2R* (Km^r/Gm^r)	This work
pJD123	pCreDelR carrying *lox3R* (Km^r/Gm^r)	This work
pJD124	pCreDelR carrying *lox4R* (Km^r/Gm^r)	This work
pJD125	pCreDelR carrying *lox5R* (Km^r/Gm^r)	This work

DNA sequences necessary for integration downstream of the *exo* gene cluster via double homologous recombination. In contrast, pK18mobsacB derivatives pJD79 and pJD80 carry a DNA sequence for integration upstream of the *exo* gene cluster followed by *loxL* and *loxLR*, respectively. For construction of *S. meliloti* strain JDSm80, SmCre was transformed with pJD35 and pJD80 including sucrose selection mediated deletion of the backbone of pJD35. JDSm84 is a derivative of SmCre which was constructed by genomic integration of pJD78 and pJD79, including sucrose selection mediated deletion of the backbone of pJD78.

pCreDelL and pCreDelR are designed for CFGE-mediated Cre/*lox* deletions. pCreDelR is a pK18mob2 derivative carrying a gentamicin resistance cassette and *loxR*. The insert serves as PCR template and is necessary for CFGE-mediated integration of *loxR* downstream of the *exo* gene cluster close to the *thiD* locus. pCreDelL is based on pK18mobsacB and comprises *loxL* and a sequence homologous to a target site upstream of the *exo* gene cluster at the *exsA* locus. Construction of JDSm85, utilizing pCreDelL and pCreDelR, is part of the method described in Fig. 4B and Table S1. Constructs used to examine spacer mutant *lox* sites are based on pCreDelL and pCreDelR. For this purpose the *loxR* sequence of pCreDelR was replaced with spacer and arm mutant *lox* sites *lox1R-lox5 R* giving rise to constructs pJD121 (*lox1R*), pJD122 (*lox2R*), pJD123 (*lox3R*), pJD124 (*lox4R*) and pJD125 (*lox5R*). Furthermore, *loxL* carried by pCreDelL was replaced with a *loxL* array, which was obtained as synthetic DNA fragment (Integrated DNA Technologies). The array

consists of six *lox* sites (*loxL* and *lox1L-lox5L*). These are separated by 25 nucleotides containing individual restriction sites and 19 bp "neutral" DNA spacers designed with R2oDNA (Casini et al., 2014). CFGE-mediated construction of respective *S. meliloti* strains gave rise to JDSm128-133 which exhibit the *loxL* array upstream of the *exo* gene cluster and one of each spacer mutant *loxR* derivatives downstream of the *exo* cluster. The detailed construction of JDSm128-133 is described in Table S1.

2.3. DNA manipulation and plasmid DNA extraction

Standard techniques were employed for common cloning procedures (Sambrook and Russel et al., 2001). For generation of *lox* sites, oligonucleotides were 5′ phosphorylated using T4 Polynucleotide Kinase (Fermentas) and subsequently hybridized. Plasmid DNA from *E. coli* strains and *S. meliloti* Rm1021 was purified using the E.Z.N.A. Plasmid Mini Kit (Omega Bio-Tek) following the supplier's instructions. For gel extractions, the illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences) was used according to manufacturer's recommendations. Dephosphorylation of 5′-ends of DNA was achieved by use of FastAP Thermosensitive Alkaline Phosphatase (Fermentas).

2.3.1. Plasmid extraction from E. coli and S. meliloti strains for assessing electroporation efficiency

Cultures of *E. coli* DH5α109-pSRKGm and DH5α strains bearing pSRKKm-*ccrM*, pJD29 and pJD30 (plus empty vectors pSRKGm/Km or pPHU231 as a control) were diluted to an OD_{600} of 0.001 and incubated as triplicates overnight in 12 ml LB medium supplemented with appropriate antibiotics, and IPTG (50, 500 μM) or arabinose (0.2%). In order to gain a homogeneous quality of DNA samples, plasmid DNA was purified from three replicates and subsequently pooled. Cultures of *S. meliloti* Rm1021 carrying pSRKGm were adjusted to an OD_{600} of 0.01 and incubated as triplicate overnight in 12 ml TY medium supplemented with streptomycin and gentamicin. Plasmid DNA was purified from two replicates and subsequently pooled. DNA quality and quantity was measured using a NanoDrop™ spectrophotometer (Thermo Fisher Scientific).

2.4. Introduction of (plasmid) DNA into S. meliloti

2.4.1. Conjugation

Recipient *S. meliloti* strains were grown overnight on TY agar plates supplemented with appropriate antibiotics. Cultures of *E. coli* S17-1 donor strains were diluted to an OD_{600} of 0.005 and 100 μl of cell suspension was plated on LB agar plates containing appropriate antibiotics. After incubation overnight at 37 °C, both *E. coli* and *S. meliloti* cultures were resuspended in 0.9% NaCl and adjusted to an OD_{600} of 2.0. 150 μl of *S. meliloti* suspension and 50 μl of *E. coli* suspension were mixed. Subsequently, 50 μl of mixed cell suspension was spotted onto a TY agar plate without antibiotics and incubated at 30 °C overnight. Cells were recovered with a sterile spatula and resuspended in 1 ml 0.9% NaCl. For determination of conjugation efficiency, the cell suspension was additionally adjusted to an OD_{600} of 3.5 and appropriate dilutions were plated on selective TY agar plates. Efficiency of conjugation was expressed as the number of colony forming units (CFU) per ml suspension.

2.4.2. Electroporation

S. meliloti strains Rm1021, SmΔhsdR, SmCreΔhsdR, *A. tumefaciens* C58 and *X. campestris* B100 were prepared for electroporation as described by Ferri et al. (2010). Electro-competent cells were thawed on ice, mixed with plasmid DNA (1–3 μl) and transferred to an electroporation cuvette (Peqlab) with an inter-electrode distance of 0.1 cm. Electroporation was performed with an Eporator (Eppendorf) at 21 kV/cm (*S. meliloti*) and 25 kV/cm (*A. tumefaciens*, *X. campestris*) with a time constant fixed at ∼5 ms. After pulse application, *S. meliloti* cells were resuspended in 1 ml TY medium and recovered at 30 °C with shaking at 200 rpm for 4 h. *A. tumefaciens* and *X. campestris* were resuspended in 1 ml LB medium and recovered at 30 °C, with shaking at 200 rpm for 3 h. Finally, appropriate dilutions were plated on selective TY/LB agar plates. For determination of electroporation efficiency of *S. meliloti*, 50 ng of pSRK derivatives, and 100 ng DNA of pK18mob2 derivatives were applied. In case of the plasmid mix purified from DH5α109/pSRKGm, 150 ng of DNA was used. Efficiency of electroporation was defined as the number of CFU per ml suspension. If necessary, an unpaired *t*-test was used to determine significance.

2.5. Cloning-free genome editing (CFGE)

The CFGE procedure is described in Section 3.2. DNA products were generated using Q5 High-Fidelity DNA Polymerase (NEB) and subsequently 5′ phosphorylated applying T4 Polynucleotide Kinase (NEB). For blunt-end ligation of appropriate DNA fragments, T4 DNA Ligase (Fermentas) and Taq DNA Ligase (NEB) were used according to manufacturer's recommendations. In this study, CFGE-based gene inactivation was based on PCR fragments comprising *aacC1* conferring gentamicin resistance (1600 ng, ∼1000 bp) and a sequence for homologous recombination with the target site (400 ng, ∼500 bp). Assembly via ligase cycling reaction (LCR) was carried out according to an optimized protocol using 3 nM of DNA parts and 30 nM of bridging oligonucleotide (Tm of half bridging oligonucleotide = 60 °C) (de Kok et al., 2014). Ligation products were purified using the E.Z.N.A. Cycle-Pure Kit (Omega Bio-Tek). Subsequently, excess water was evaporated in order to gain appropriate sample volumes for electroporation. Regarding LCR products, two samples were pooled and entirely used for transformation. Detailed information about construction of CFGE fragments used in this study is given in Table S1.

2.6. Cre-mediated deletion of the exo gene cluster and examination of spacer mutant lox sites

Cre/*lox* applications are based on the taurine-inducible *cre* expression system described by Harrison et al. (2011). In this study, CFGE-mediated Cre/*lox* experiments are based on derivatives of strains SmCre and SmCreΔhsdR. For *cre* induction, the respective strains were inoculated in TY medium supplemented with 10 mM taurine and 5 μg/ml Tc. Cell density was adjusted to an OD_{600} of 0.01. After incubation at 30 °C with shaking at 200 rpm for 48 h, dilution series were plated on LB medium containing Str and 10% sucrose. In order to determine the number of CFU/ml after *cre* induction, cell suspensions were adjusted to an OD_{600} of 1. Since tetracycline occasionally impairs cell growth, thereby effecting efficiency of Cre-mediated recombination, strains JDSm80, JDSm84, and JDSm128-133 were inoculated in TY medium supplemented with 10 mM taurine and 600 μg/ml Str and incubation after *cre* induction was shortened to 24 h.

3. Results and discussion

3.1. Transformation of S. meliloti

The standard method for transformation of *S. meliloti* is *E. coli*-mediated conjugal transfer of plasmids bearing an RK2/RP4 origin of transfer (*oriT*). The system is based on application of a helper plasmid providing the appropriate conjugal transfer functions (Figurski and Helinski, 1979), or on *E. coli* strains S17-1 or SM-10 which carry chromosomally integrated conjugal transfer functions derived from RP4 (Simon et al., 1983). Both approaches

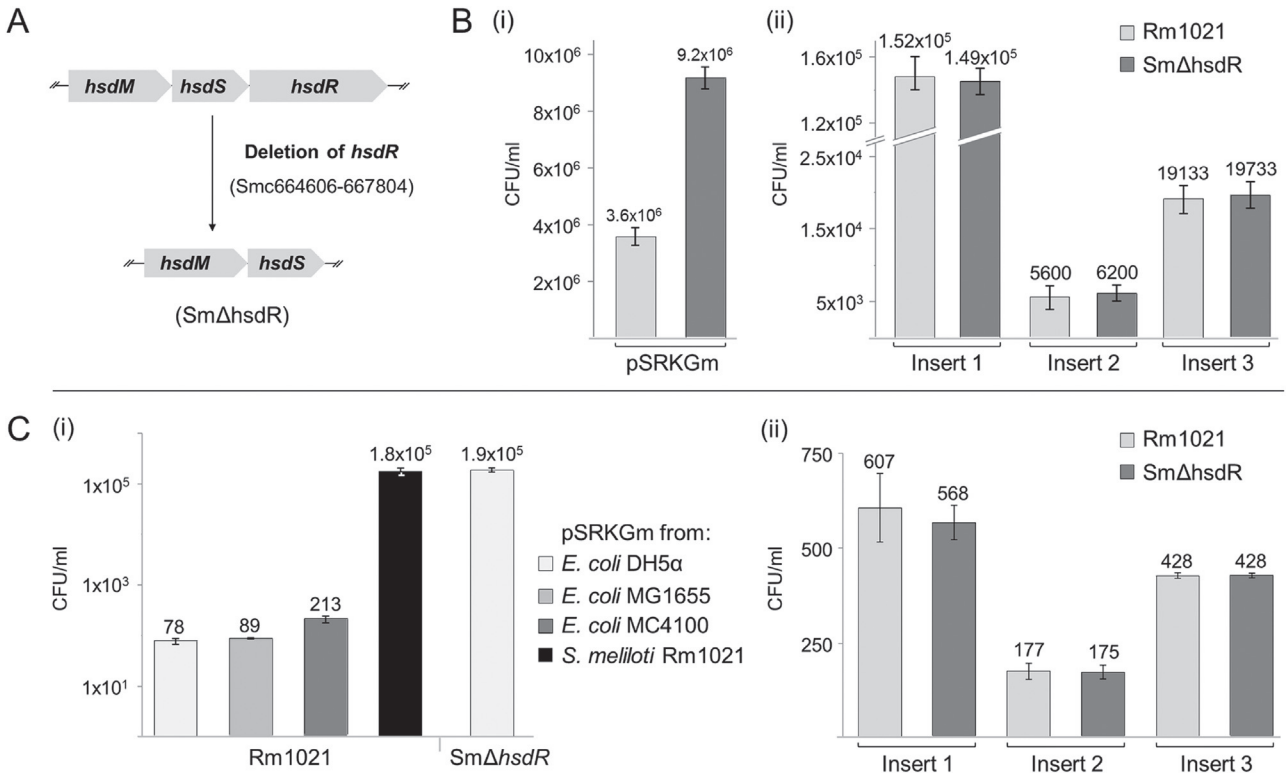


Fig. 1. *S. meliloti* strain SmΔhsdR shows enhanced transformation efficiency. (A) Scheme of the *hsdR* deletion site on the chromosome. The promoter and coding sequence of *hsdMS* remained unaffected. (B) Transformation rates of SmΔhsdR in comparison to Rm1021 after conjugation of replicative (i) or integrative (ii) constructs. Integrative constructs (Insert 1–3) carry ~550 bp fragments homologous to *exoP* (pJD23), *exoI* (pJD24), and *exsA* (pJD25). (C) Transformation rates of SmΔhsdR in comparison to Rm1021 after electroporation of replicative (i) or integrative (ii) constructs. The replicative plasmid pSRKGm was purified from different *E. coli* strains and *S. meliloti* Rm1021. Integrative constructs (Inserts 1, 2 and 3) correspond to pJD23, pJD24, and pJD25, respectively, and were purified from *E. coli* DH5α. For electroporation 50 ng and 100 ng of pSRKGm and pK18mob2 derivatives, respectively, were applied.

have some drawbacks, such as co-transfer of the helper plasmid or chromosomal DNA of the *E. coli* donor into the *S. meliloti* recipient strain. The latter is occasionally caused by a fully functional Mu temperate prophage, present in the genomes of *E. coli* S17-1 and Sm-10, which is potentially mobilized and transferred into the recipient strains (Babic et al., 2008; Ferrières et al., 2010). Even plasmids transferred from *E. coli* S17-1 to other bacterial species often become modified by insertion of chromosomal DNA derived from the donor (Strand et al., 2014). A plasmid-based mobilization system circumvents *E. coli* S17-1/SM10-related problems (Strand et al., 2014). Nevertheless, potential incompatibility between the (pBBR1-based) helper plasmid and the vector which is intended to be transferred still restrains applicability. An alternative is introduction of DNA via electroporation. However, transformation by electroporation of wild type *S. meliloti* suffers from low efficiency.

3.1.1. *Deletion of hsdR results in an increase of* S. meliloti *transformation efficiency*

A previous study identified the restriction endonuclease HsdR to be responsible for the low electroporation capability of the *S. meliloti* Rm1021 wild type strain (Ferri et al., 2010). Compared to the wild type, a mutant strain carrying a plasmid integration in the chromosomal locus of *hsdR* showed a ~20–200-fold increase in transformation efficiency using plasmid DNA extracted from different rhizobia for electroporation, whereas a 100-fold increase was reported employing plasmid DNA purified from *E. coli* DH5α (Ferri et al., 2010). To take advantage of this increased transformation efficiency, we constructed the stable *S. meliloti* knockdown strain SmΔhsdR by deletion of the complete *hsdR* coding sequence (Fig. 1A) and determined its transformation rate applying electroporation and *E. coli* S17-1-mediated conjugation (Fig. 1B, C). As previously reported, deletion of *hsdR* resulted in an increase in transformation efficiency by the replicative plasmid pSRKGm. In accordance with Ferri et al. (2010), conjugal transfer of this plasmid to SmΔhsdR generated ~2.6-fold more transformants than applying Rm1021 as recipient (Fig. 1B (i)). Electroporation of Rm1021 with plasmid DNA purified from *E. coli* strains DH5α, MC4100 and MG1655 commonly resulted in poor transformation rates of about 80–200 CFU/ml (Fig. 1C (i)). In contrast, electroporation of SmΔhsdR strikingly enhanced the transformation rate about 2400-fold compared to the wild type recipient and reached $1.9*10^5$ CFU/ml. This increase was much higher than expected from the study of Ferri et al. (2010) and matched the transformation rate achieved with self-derived plasmid DNA introduced to Rm1021 by electroporation. In order to assess the transformation efficiency by plasmids non-replicative in *S. meliloti*, which require integration into the genome for maintenance, different homologous regions for integration into pSymB were inserted into suicide vector pK18mob2. Irrespective of the mode of transfer, transformation rates by these integrative plasmids were not affected by deletion of *hsdR* (Fig. 1B, C (ii)). Thus, genomic integration of plasmid DNA is most likely mainly limited by the frequency of homologous recombination events and not by HsdR-mediated degradation of introduced DNA.

In conclusion, SmΔhsdR offers strong benefits for transfer of replicative plasmids by electroporation. Transformation of *S. meliloti* Rm1021 by pSRKGm applying electroporation was about 40.000-fold less effective than conjugation. Use of SmΔhsdR reduced this difference to about 20-fold (Fig. 1C (i)). Nonetheless, the transformation rate of the *S. meliloti* wild type by common constructs introduced via electroporation might still be sufficient. However, increasing sizes of DNA constructs to be established in

the recipient in synthetic biology approaches require highly efficient procedures. Accordingly, electroporation-mediated transfer of the 19 kb replicative vector pPHU231 to *S. meliloti*, employing ~200 ng vector DNA purified from *E. coli* DH5α, was unsuccessful unless SmΔhsdR was the recipient (<3 CFU/ml were obtained).

3.1.2. *Increasing the efficiency of electroporation-mediated transformation of the* S. meliloti *wild type*

Here we describe another way to enhance electroporation-mediated transformation efficiency of *S. meliloti*, in case the wild type genetic background has to be retained. The negative effect of *hsdR* on the transformation efficiency and the increased transformation rates using self-derived plasmid DNA (Fig. 1C) indicate discrimination of DNA bearing a "foreign" methylation pattern. Therefore, we placed the DNA methyltransferase-encoding genes *ccrM*, *hsdM* and *hsdM-hsdS* from *S. meliloti* under the control of an IPTG-inducible *lac* promoter in vectors pSRKKm or pSRKGm and introduced these constructs to *E. coli* DH5α. Prior to plasmid DNA purification and electroporation of the *S. meliloti* wild type, expression of these DNA-methyltransferase genes was induced. The resulting transformation rates are shown in Fig. 2A. In comparison to an empty vector control, full induction of *ccrM* expression resulted in a low but statistically significant increase in the transformation rate. This was surprising since CcrM is supposed to be a component of the cell cycle control machinery (Wright et al., 1997). A role for CcrM-mediated DNA methylation in protection from restriction enzyme cleavage has so far not been discussed. *hsdM* and *hsdM-hsdS* affected the transformation rate of *S. meliloti* even when expression was not induced, probably due to leaky expression. Whereas *hsdM* completely abolished transformation, full induction of *hsdM-hsdS* expression resulted in a 10-fold increase.

In these experiments an additional plasmid-borne ectopic copy of *hsdM-hsdS* was introduced by the electroporation to the *S. meliloti* recipient. To rule out an effect of this additional *hsdM-hsdS* copy on the transformation rate the empty vector pSRKKm was introduced into the plasmid donor strains *E. coli* DH5α and DH5α containing the pSRKGm-derivative pJD30 driving expression of *hsdM-hsdS*. pSRKKm purified from these strains by gel extraction was used for electroporation of the *S. meliloti* wild type resulting in a ~85-fold higher transformation rate when the plasmid was re-isolated from the fully induced *hsdM-hsdS* expressing strain (Fig. S1A). This suggests that propagation of plasmid DNA in *E. coli* expressing the *S. meliloti hsdM-hsdS* genes probably results in DNA methylation that increases stability after transfer to the DNA restriction-modification potent *S. meliloti* wild type.

Based on this observation a pPHU231-based low copy helper plasmid carrying *hsdM-hsdS-ccrM* under the control of both PBAD and the plasmid-derived *lac* promoter was generated and introduced to *E. coli* DH5α. The resulting strain DH5α109 was co-transformed with pSRKGm (Fig. S1B) and expression of the DNA methyltransferase genes was induced overnight. Electroporation of the *S. meliloti* wild type with purified plasmid DNA yielded a ~10-fold increase in the transformation rates in comparison to the control even when *hsdM-hsdS-ccrM* expression was not induced (Fig. 2B). Surprisingly, induction with 0.2% arabinose failed to increase transformation rates compared to the control. A similar effect was caused by induction of *hsdM-hsdS-ccrM* expression by 0.5 M IPTG (data not shown), implying that expression levels may be critical. Co-transformation of *S. meliloti* with pSRKGm and the helper plasmid was not observed. This indicates that gel-extraction of the plasmid to be transferred to *S. meliloti* is not necessary because the low copy number of the helper plasmid does most likely not allow effective transfer by electroporation.

We noted that growth conditions of the *E. coli* strain carrying the helper plasmid and the plasmid to be propagated significantly influence the transformation rate of *S. meliloti* electroporated with

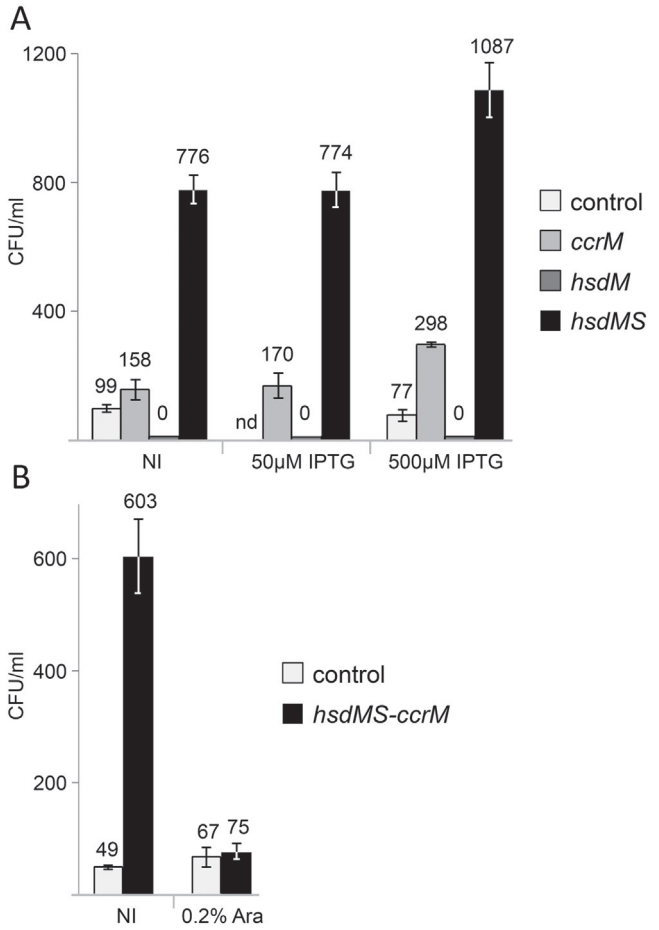


Fig. 2. Expression of DNA methyltransferase genes in the *E. coli* plasmid donor strain affects electroporation-mediated transformation rates of *S. meliloti*. (A) *E. coli* DH5α strains bearing either empty pSRKGm (control) or pSRK carrying the *S. meliloti* DNA methyltransferase genes *ccrM* (pSRKKm:*ccrM*), *hsdM* (pJD29) and *hsdM-hsdS* (pJD30) were used as plasmid donor strains. Plasmid DNA was purified from these strains after incubation in LB medium supplemented with 0, 50 and 500 µM IPTG. For electroporation of *S. meliloti* Rm1021, 50 ng plasmid DNA were applied. The control was not performed with 50 µM IPTG (nd). NI, not induced. (B) *E. coli* strains DH5α109 (DH5α carrying the helper plasmid pPHU231 with *hsdM-hsdS-ccrM*) and DH5α/pPHU231 (serving as control) were used as pSRKGm donor strains. Plasmid DNA was purified from these strains after incubation in LB medium supplemented with 0 and 0.2% arabinose. For electroporation of *S. meliloti* Rm1021 150 ng of plasmid DNA was used. NI, not induced.

the isolated plasmid. This offers a significant potential for optimization which will benefit from a functional elucidation of the heterologously expressed DNA methyltransferase genes in *E. coli*. Nevertheless, we provide a helper plasmid-based system which improves the transformation rate of *S. meliloti* electroporated with plasmid DNA purified from *E. coli*. This is a useful step forward in improving electroporation-mediated transformation of the *S. meliloti* wild type and offers an alternative to more elaborate *in-vitro* methylation methods (Chen et al., 2008; Donahue et al., 2000).

3.2. *method for cloning-free genome editing (CFGE) facilitates genome engineering in* S. meliloti

Integration of DNA fragments into the genome of *S. meliloti* commonly comprises several working steps. After design and cloning of the respective constructs in *E. coli*, plasmid DNA is introduced into *S. meliloti* via conjugation with the aid of *E. coli* helper strains. A single homologous recombination event leads to integration of both the sequence of interest and the backbone of the plasmid, which is non-replicative in *S. meliloti*. Consequently, transconjugants cannot

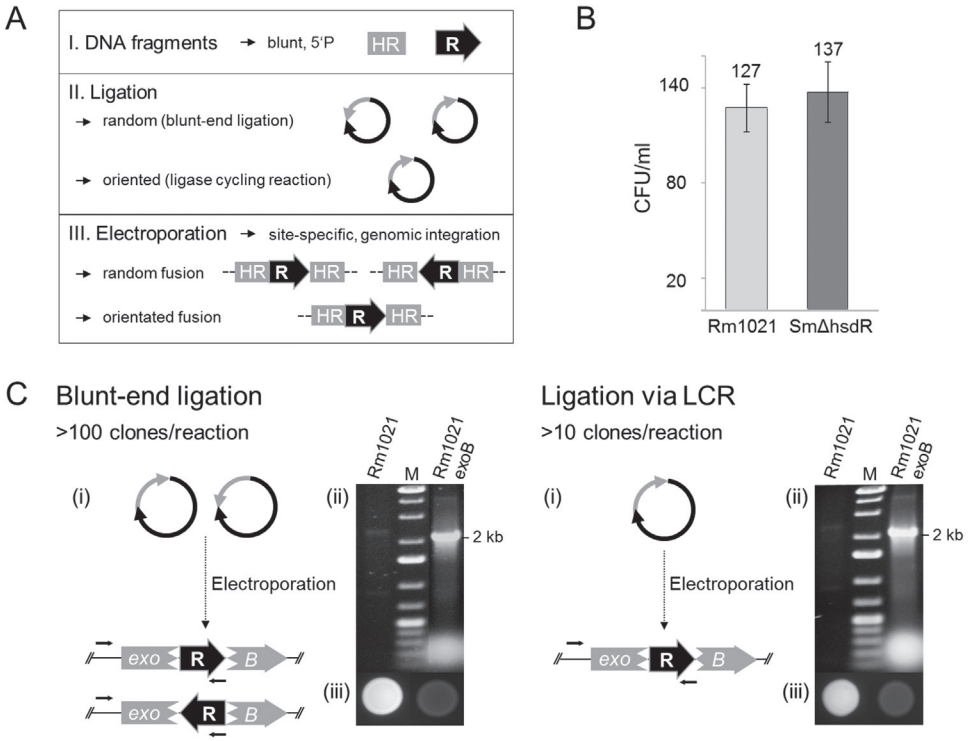


Fig. 3. Cloning-free genome editing (CFGE) facilitates gene knockouts in *S. meliloti*. (A) Scheme of a CFGE-mediated genomic integration in three steps. The minimal setup comprises a region homologous to the target site (HR) and a resistance cassette (R). Fragments were blunt-ended and 5′ phosphorylated (5′P). DNA fragments were circularized either by common blunt-end ligation or ligase cycling reaction (LCR). *S. meliloti* cells were electroporated with purified DNA. (B) Electroporation-mediated transformation rate of *S. meliloti* SmΔhsdR in comparison to Rm1021 applying CFGE for interruption of *exoB*. Electrocompetent cells were transformed with 450 ng of purified ligation products. (C) CFGE-mediated inactivation of *exoB* resulted in generation of strain Rm1021 exoB (for details see Table S1). (i) *S. meliloti* was electroporated with circular DNA fragments comprising a homologous region (grey circle) and a gentamicin resistance cassette (black circle). Arrows indicate primers used for PCR. (ii) PCR-amplification of the insertion site using the indicated primers resulted in ~2 kb DNA fragments when *aacC1* was successfully integrated. (iii) Calcofluor fluorescence of PCR-verified *exoB* knockout mutants in comparison to the wild type.

be simply transformed with further integrative constructs based on the same or related vectors, since homologous parts of the backbones, such as the origin of transfer, cause unintended co-integration. In order to eliminate such homologous regions, usually *sacB* encoding the levansucrase is used as conditional lethal marker to enable sucrose selection for loss of the plasmid backbone due to a second homologous recombination event in a time-consuming downstream screening procedure (Schäfer et al., 1994).

To expedite engineering of the *S. meliloti* genome, we developed a method for cloning-free genome editing (CFGE). Genomic integration via transformation by linear DNA molecules is a well established method in model organisms such as *E. coli* or *B. subtilis* (Niaudet et al., 1985; Zhang et al., 1998). Electroporation of *S. meliloti* with linear DNA fragments aiming at integration via homologous recombination was not successful. However, genomic integration of self-ligated DNA fragments introduced to *S. meliloti* by electroporation was achieved. Fig. 3A shows the minimal setup of CFGE by a circular product comprising a homologous region and a selection marker.

The suggested DNA ligation methods require fragments with 5′ phosphorylated blunt ends which can be generated with appropriate DNA polymerases. Subsequently, purified DNA is fused via blunt-end ligation. Depending on the needs of the application, two approaches proved to be practical. When a defined orientation is not required and potential concatemers need not be completely excluded, common blunt-end ligation is sufficient. If directed fusion of (several) fragments is desired, use of an *in-vitro* assembly method, such as the ligase cycling reaction (LCR), is recommended. Finally, electrocompetent *S. meliloti* cells are transformed with the purified ligation mix. Transformants that are resistant to the respective antibiotic are likely to carry the fragment(s) of interest at the intended target site. To date, this is the fastest genomic integration method available for *S. meliloti*.

3.2.1. CFGE enables gene knockouts executed in one day

CFGE can effectively be used for gene inactivation. Fig. 3C shows the procedure for inactivation of *exoB* on the pSymB megaplasmid. Both a gentamicin resistance cassette (*aacC1*) and a 500 bp DNA fragment homologous to the coding sequence of *exoB* were PCR amplified and either fused via blunt-end ligation or LCR. Electroporation of *S. meliloti* with the purified ligation mix resulted in gentamicin resistant strain Rm1021 exoB. Both approaches produced a sufficient number of colonies.

However, different yields are accounted for by different amounts of DNA used for ligation (see materials and methods). Integration of *aacC1* into *exoB* was validated via PCR. Over 95% of tested clones were positive. Whereas blunt-end ligation led to both a forwardly and reversely integrated copy of *aacC1*, LCR-based CFGE exclusively generated clones bearing a forwardly integrated resistance gene (Fig. S2). Since *exoB* mutants are defective in exopolysaccharide production due to loss of UDP-glucose 4-epimerase activity (Leigh et al., 1987), PCR-verified clones were further confirmed in a calcofluor fluorescence assay (Fig. 3C, Fig. S2). As previously found for genomic plasmid integration (Fig. 1C), the electroporation-mediated transformation rate by circular DNA fragments was not affected by deletion of *hsdR*, but appeared to be limited by other factors (Fig. 3B).

We also successfully applied this method in *A. tumefaciens* and *X. campestris* to knockout *exoB* and *gumD*, respectively (Fig. S4). In conclusion, CFGE appears to be a suitable method for rapid gene

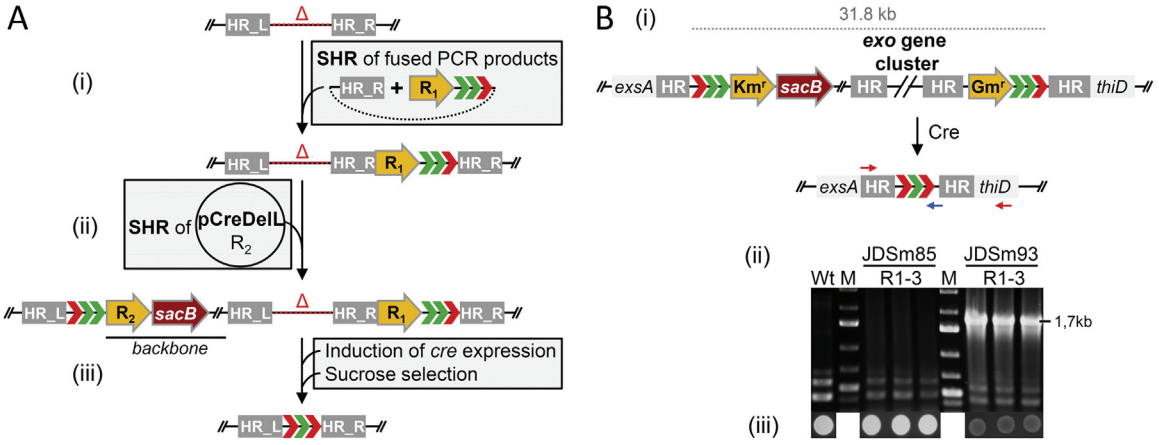


Fig. 4. (A) Scheme of a Cre/*lox* system for serial deletions of large DNA regions (Δ). (i) *loxR* (arrows, red indicates arm mutation) is inserted via CFGE. A sequence homologous to a target site downstream of region Δ (HR_R) is PCR-amplified from genomic DNA. The fragment comprising a resistance cassette (R_1) and *loxR* is PCR-amplified from vector pCreDelR. (ii) pCreDelL is a pK18mobsacB derivative carrying a sequence homologous to a target site upstream of region Δ (HR_L). (iii) Cre-mediated recombination results in deletion of previously inserted sequences and region Δ. As a result from the non-homologous recombination event, left/right arm mutated *lox* sites convert into *loxLR*, which remains in the genome. SHR, single homologous recombination event. (B) Cre-mediated deletion of a ~32 kb region (SMb1166733-1190671) including the ~23 kb *exo* gene cluster from strain JDSm85. (i) Integration of *loxR* by CFGE (see Table S1) and subsequent conjugation of pCreDelL carrying *loxL*. Cre-mediated recombination resulted in JDSm93. Km^r/Gm^r: kanamycin/gentamicin resistance cassette. Red arrows: Primers used for PCR. Blue arrow: sequencing primer. (ii) A region spanning the deletion site was PCR-amplified with primers binding close to *exsA* and *thiD*, resulting in ~1700 bp PCR products. Wt: *S. meliloti* Rm1021. R1-3: representative replicates. M: GeneRuler 1 kb Plus (Thermo Fisher Scientific). (iii) Calcofluor fluorescence of PCR-verified mutants in comparison to JDSm85 and wild type *S. meliloti*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

inactivation in representatives of the α- and γ-proteobacteria. No plasmid derived sequences are integrated into the genome, and reapplication is assumed to be exclusively limited by the number of available resistance markers. In future studies, inclusion of the *sacB* gene and a second homologous region in the circular construct will be tested aiming at application of CFGE for marker-free gene deletion. Since integrations by a single-crossover also bear the potential to revert to wild type, this approach may allow for releasing the resistance cassette and generation of stable integrations or deletions.

3.2.2. An enhanced Cre/lox system facilitates comprehensive deletion studies

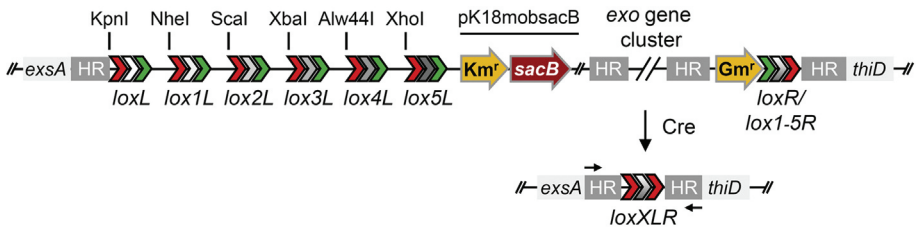
Based on a taurine inducible *cre* expression system in *S. meliloti* (Harrison et al., 2011), we constructed the plasmid backbone-free derivatives SmCre and SmCreΔhsdR as stable *S. meliloti* platform strains for Cre-mediated manipulation of large genomic regions. The SmCreΔhsdR strain does not confer specific advantages in Cre/*lox* procedures but might prove beneficial to downstream applications. Fig. 4A depicts the underlying procedure for deletion of large DNA fragments. Sites flanking the target region to be deleted have to be equipped with Cre recognition (*lox*) sites mediating non-homologous recombination and *sacB* for positive selection on loss of the target region. Since a *loxP* site remaining in the genome as result of the non-homologous recombination event would interfere with subsequent Cre-mediated reactions, left arm and right arm mutant *loxL* and *loxR* sites were used (corresponding to lox71 and lox66; Albert et al., 1995). In a first step, CFGE allows for integration of *loxR* into the right flank without any cloning effort (Fig. 4A (i)). A single homologous recombination event is sufficient for each step, since in the second step, pCreDelL, a pK18mobsacB derivative carrying *loxL* and a homologous region for integration into the left flank, can be introduced by conjugation or electroporation without risking co-integration (Fig. 4A (ii)). Finally, Cre-mediated recombination between head-to-tail oriented *lox* sites results in deletion of the target region and previously inserted sequences (Fig. 4A (iii)). For deletion of the whole ~23 kb *exo* gene cluster directing biosynthesis of the acidic exopolysaccharide succinoglycan (Becker et al., 1993; Reuber and Walker, 1993), strain JDSm85 was prepared as described above, followed by induced *cre* expression and sucrose selection on the loss of the conditional lethal marker *sacB* (Fig. 4B (i)). 67% of the sucrose resistant, kanamycin sensitive transconjugants obtained from three independent approaches were tested positive for the deletion by a PCR assay (Fig. 4B (ii), Fig. S3A). Calcofluor fluorescence assays for succinoglycan production and sequencing of a DNA fragment spanning the deletion site further confirmed loss of the *exo* genes (Fig. 4B, Fig. S3A).

The recombination efficiency was higher (up to 100%, data not shown) in a similar approach using native *loxP* sites instead of *loxL/R*. Nevertheless, usage of these left arm and right arm *lox* mutants for Cre-mediated site-specific recombination, in combination with sucrose selection on loss of *sacB*, did not require extensive screening efforts to identify positive clones, despite of the large size of the deleted region. As expected, Cre-mediated recombination between *loxL* and *loxR* resulted in a *loxLR* site. This method allows for efficient marker-free deletion, even of large DNA regions, and offers the possibility to insert sequences for downstream applications at the deletion site. A 34 bp *loxLR* site is the minimal scar remaining in the genome. However, a fully mutated *loxLR* site is not supposed to effectively interfere with subsequent Cre/*lox* applications (Fig. S3B). Moreover, employing CFGE for genomic insertions of *lox* sites (and markers) offers substantial time savings in genome engineering procedures.

3.2.3. Examination of mutant spacer-containing lox sites in S. meliloti

In order to extend the set of non-promiscuous *lox* sites in *S. meliloti* we chose five *lox* spacer derivatives which were previously described in the literature as self-recombining and non-promiscuous, at least with the wild type *loxP* site (Fig. S5A). *lox1* and *lox2* correspond to *lox* sites m3 and m7 (Langer et al., 2002) which were shown to mediate non-promiscuous Cre-recombination in HEK293 cells. *lox3* corresponds to *lox2272* (Lee & Saito; 1998) which was previously reported to specifically self-recombine in an *in-vitro* system using Cre heterologously produced in mammalian cells. In the course of a high-throughput *in-vitro* assay Missirlis et al. (2006) identified two more spacer mutants (named *lox4* and *lox5* in this

A



B

lox site spacer origin	*loxP*	*lox1*	*lox2*	*lox3*	*lox4*	*lox5*
CFU after Cre induction	$3.3x10^8$	$1.8x10^7$	$2.6x10^7$	$1.5x10^8$	$2.6x10^7$	$5.1x10^5$
(STDEV)	($\pm1.7x10^8$)	($\pm1.6*10^7$)	($\pm2.4*10^7$)	($\pm1.3*10^8$)	($\pm2.5*10^7$)	($\pm3.9*10^5$)
Km resistant clones	0%	0%	0%	0%	0%	3%
Positive in PCR	100%	98%	100%	100%	97%	7%
and restriction digestion	100%	100%	100%	100%	100%	100%

Fig. 5. L/R-spacer mutant *lox* site assay in *S. meliloti*. (A) Schematic representation of the genetic setup. In order to examine recombination efficiency and specificity of left/right arm mutated Cre recognition sites carrying mutated spacer regions (*lox1L/R-lox5L/R*) or the native spacer (*loxL/R*), a *lox* site array was integrated upstream of the ~23 kb *exo* gene cluster on pSymB. Furthermore, individual *lox* sites (*loxR* and *lox1R-lox5R*) were integrated downstream of this gene cluster. After induction of *cre* expression, cultures were plated on sucrose-containing LB agar plates for selection on Cre-mediated recombination events. Successful deletion of the *exo* gene cluster was verified via PCR using oligonucleotides specifically binding to regions up- and downstream of this gene cluster (black arrows). Cre-mediated recombination was expected to result in conversion of *loxL/R* derivatives into a fully mutated left and right arm mutant *lox* site (*loxXLR*). Due to individual restriction sites flanking the *loxL* derivatives, restriction analysis of PCR products comprising the deletion site allowed for detection of specific recombination events. (B) Analysis of Cre/*lox* recombination mediated by spacer derivatives derived from *lox1-lox5* (JDSm128-132) and *loxP* (JDSm133). After Cre induction *S. meliloti* cultures of each strain were normalized to an OD_{600} of 1 and incubated on LB agar supplemented with 10% sucrose. After three days of incubation at 30 °C the number of CFU per ml was determined. Independent clones were screened for loss of kanamycin resistance (n = 50) and positive (kanamycin sensitive) clones were tested for loss of the *exo* gene cluster via PCR with primers binding up- and downstream of the *exo* gene cluster (n = 30). Proper deletion of the *exo* gene cluster was confirmed for three positive clones by sequencing a PCR-amplified fragment containing the deletion site. Moreover, PCR-verified clones were tested for specific recombination between matching *lox* derivatives by digestion with appropriate restriction enzymes flanking the respective *loxL* site (n = 20 (*lox2*, *lox3*, *loxP*)/19 (*lox1*)/18 (*lox4*)/1 (*lox5*)). STDEV, standard deviation.

study) which also hold promise for highly specific recombination events. Recombination efficiency between these *lox* sites was tested employing an array of *loxL* sites upstream of the *exo* gene cluster and one of the *loxR* sites downstream of this gene region.

pCreDelR and its derivatives pJD121-125 served as template for PCR amplification of *aacC1-loxXR* fragments which were integrated downstream of the *exo* genes by applying CFGE. Subsequently, the pCreDelL derivative pJD116 carrying a *lox* site array was introduced by conjugation and integrated upstream of the *exo* gene cluster. The resulting strains JDSm128-132 (*lox1R-lox5R*) and JDSm133 (*loxR*) were subjected to Cre-mediated deletion of the *exo* gene cluster (Fig. 5A). The arrayed *loxL* sites were separated by 19 bp linker sequences which contain unique recognition sites for several restriction enzymes (Fig. S5A). This allowed examination of recombination products by characterization of restriction patterns of PCR products spanning the deletion site. All tested mutant spacer-containing *lox* sites mediated deletion of the ~23 kb gene cluster in a highly specific manner (Fig. 5B).

Whereas JDSm133 (*loxL* array <> *loxR*) and JDSm130 (*loxL* array <> *lox3R*) exhibited the highest Cre/*lox* recombination efficiencies (3.3×10^8 and 1.5×10^8 CFU/ml, respectively), other strains yielded ~10-fold (JDSm128, JDSm129 and JDSm131) or even ~650-fold (JDSm132) less CFU/ml compared to JDSm133. Except for JDSm132 (*loxL* array<> *lox5R*) all tested clones lost kanamycin resistance due to *sacB* counter-selection. 97–100% of kanamycin sensitive clones derived from JDSm128-131 and JDSm133, but only 7% of kanamycin sensitive clones derived from JDSm132 (*lox* array <> *lox5R*), were tested positive for deletion of the *exo* gene cluster via PCR (Fig. 5B and S5B). Digestion with restriction enzymes cleaving up- and downstream of each *lox* site derivative of amplicons spanning the deletion site of PCR-validated clones confirmed highly specific Cre-mediated recombination. Recombination between different *lox* spacer derivatives was not detected. In conclusion, *lox1-lox5* successfully mediated non-promiscuous Cre/*lox* reactions in *S. meliloti* and proved to mediate recombination as reliably as the native spacer. The *lox5* site exhibited a very low recombination efficiency. In contrast, Cre recognition sites based on the native or *lox3* spacer seem to be the most favored substrates for the recombinase. Nonetheless, *lox1*, *lox2* and *lox4* also mediated sufficient recombination frequencies allowing large-scale deletions in *S. meliloti*.

4. Conclusions

Despite the enormous biotechnological, medical, and environmental importance of α-proteobacteria, the genetic toolboxes for these organisms are still mostly based on conjugation- or electroporation-based introduction of plasmid-borne DNA propagated in *E. coli*. Moreover, handy genetic systems for serial genome reorganization are lacking. On the other hand, synthetic biology approaches strive for increasingly large-scale genomic manipulations to address fundamental questions of bacterial genome organization and to provide new opportunities for molecular biotechnology. Here, we focused on advancing the genetic toolbox for *S. meliloti* serving as model system for studies of symbiotic nitrogen fixation, which is a highly important process in sustainable agriculture.

We followed up the observation of Ferri et al. (2010) that integration of a plasmid into the restriction endonuclease-encoding gene *hsdR* results in an increase in transformation rate and constructed the stable *S. meliloti hsdR* deletion mutant SmΔhsdR. This allowed for a thorough characterization of the transformation properties of this mutant which serves as platform strain for future genetic studies. In accordance with the previous study (Ferri et al., 2010), SmΔhsdR clearly showed an enhanced efficiency in transformation with replicative plasmids propagated in *E. coli*. We further introduced expression of *S. meliloti*-derived DNA methyltransferase genes in *E. coli* propagating plasmid DNA as an alternative way of weakening the transformation barrier constituted by an R-M system in *S. meliloti*.

Aiming at accelerating genome editing in *S. meliloti*, we devised a novel CFGE strategy to obviate the need for cloning genetic constructs in *E. coli* for transformation of the recipient. This approach takes advantage of the increasing number of potent *in-vitro* assembly methods and a progressive drop in DNA synthesis costs. It offers the possibility of *in-vitro* assembly of standardized pre-manufactured genetic parts with custom synthesized or PCR-amplified sequences that target sites in the genome for homologous or non-homologous recombination. Pre-manufactured genetic parts may be resistance markers, reporter genes or genetic control elements, such as promoters and terminators. Here, we demonstrated the utility of CFGE for rapid genomic integrations in *S. meliloti* and *A. tumefaciens* (both α-proteobacteria), and *X. campestris* (a γ-proteobacterium), as well as for application of the Cre/*lox* recombination system in *S. meliloti*. Our results suggest that this method is probably applicable in a much broader range of bacteria.

The Cre/*lox* recombination system combined with CFGE has several potential uses. Here we demonstrated effective and reliable deletion of a large gene cluster. But this recombination system can also generate inversions and translocations of DNA regions, and replicon fusions. In view of the availability of a set of effective and non-promiscuous *lox* site derivatives, this makes it a powerful tool for tailored genome rearrangements involving serial steps of genome manipulation.

Acknowledgements

This work was supported by the LOEWE Program of the State of Hesse (SYNMIKRO), CRC 987 (German Research Foundation), and grant 031L0010B in the framework of the ERASynBio program (BMBF, Germany). We thank Benjamin Frage and Franz Narberhaus for providing pSRKKm-*ccrM* and *A. tumefaciens* C58, respectively.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbiotec.2016.06.033.

References

Abremski, K., Hoess, R., 1984. Bacteriophage P1 site-specific recombination: purification and properties of the Cre recombinase protein. J. Biol. Chem. 259, 1509–1514.

Albert, H., Dale, E.C., Lee, E., Ow, D.W., 1995. Site-specific integration of DNA into wild-type and mutant *lox* sites placed in the plant genome. Plant J. 7, 649–659.

Araki, K., Araki, M., Yamamura, K., 1997. Targeted integration of DNA using mutant *lox* sites in embryonic stem cells. Nucleic Acids Res. 25, 868–872.

Babic, A., Guérout, A.M., Mazel, D., 2008. Construction of an improved RP4 (RK2)-based conjugative system. Res. Microbiol. 159, 545–549.

Becker, A., Kleickmann, A., Keller, M., Arnold, W., Pühler, A., 1993. Identification and analysis of the *Rhizobium meliloti exoAMONP* genes involved in exopolysaccharide biosynthesis and mapping of promoters located on the *exoHKLAMONP* fragment. Mol. Gen. Genet. 241, 367–379.

Beringer, J.E., 1974. R factor transfer in *Rhizobium leguminosarum*. J. Gen. Microbiol. 84, 188–198.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y., 1997. The complete genome sequence of *Escherichia coli* K-12. Science 277, 1453–1462.

Capela, D., Barloy-Hubler, F., Gouzy, J., Bothe, G., Ampe, F., Batut, J., Boistard, P., Becker, A., Boutry, M., Cadieu, E., Dréano, S., Gloux, S., Godrie, T., Goffeau, A., Kahn, D., Kiss, E., Lelaure, V., Masuy, D., Pohl, T., Portetelle, D., Pühler, A., Purnelle, B., Ramsperger, U., Renard, C., Thébault, P., Vandenbol, M., Weidner, S., Galibert, F., 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. Proc. Natl. Acad. Sci. U. S. A. 98, 9877–9882.

Casini, A., Christodoulou, G., Freemont, P.S., Baldwin, G.S., Ellis, T., MacDonald, J.T., 2014. R2oDNA designer: computational design of biologically neutral synthetic DNA sequences. ACS Synth. Biol. 3, 525–528.

Chen, Q., Fischer, J.R., Benoit, V.M., Dufour, N.P., Youderian, P., Leong, J.M., 2008. *In vitro* CpG methylation increases the transformation efficiency of *Borrelia burgdorferi* strains harboring the endogenous linear plasmid lp56. J. Bacteriol. 190, 7885–7891.

de Kok, S., Stanton, L.H., Slaby, T., Durot, M., Holmes, V.F., Patel, K.G., Platt, D., Shapland, E.B., Serber, Z., Dean, J., Newman, J.D., Chandran, S.S., 2014. Rapid and reliable DNA assembly via ligase cycling reaction. ACS Synth. Biol. 3, 97–106.

diCenzo, G.C., MacLean, A.M., Milunovic, B., Golding, G.B., Finan, T.M., 2014. Examination of prokaryotic multipartite genome evolution through experimental genome reduction. PLoS Genet. 10, e1004742.

Donahue, J.P., Israel, D.A., Peek, R.M., Blaser, M.J., Miller, G.G., 2000. Overcoming the restriction barrier to plasmid transformation of *Helicobacter pylori*. Mol. Microbiol. 37, 1066–1074.

Ferri, L., Gori, A., Biondi, E.G., Mengoni, A., Bazzicalupo, M., 2010. Plasmid electroporation of *Sinorhizobium* strains: the role of the restriction gene *hsdR* in type strain Rm1021. Plasmid 63, 128–135.

Ferrières, L., Hémery, G., Nham, T., Guérout, A.M., Mazel, D., Beloin, C., Ghigo, J.M., 2010. Silent mischief: bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. J. Bacteriol. 192, 6418–6427.

Figurski, D.H., Helinski, D.R., 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. Proc. Natl. Acad. Sci. U. S. A. 76, 1648–1652.

Galardini, M., Mengoni, A., Brilli, M., Pini, F., Fioravanti, A., Lucas, S., Lapidus, A., Cheng, J.F., Goodwin, L., Pitluck, S., Land, M., Hauser, L., Woyke, T., Mikhailova, N., Ivanova, N., Daligault, H., Bruce, D., Detter, C., Tapia, R., Han, C., Teshima, H., Mocali, S., Bazzicalupo, M., Biondi, E.G., 2011. Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. BMC Genomics 12, 235.

Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M.J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R.W., Dreano, S., Federspiel, N.A., Fisher, R.F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R.W., Jones, T., Kahn, D., Kahn, M.L., Kalman, S., Keating, D.H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M.C., Pohl, T.M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorholter, F.J., Weidner, S., Wells, D.H., Wong, K., Yeh, K.C., Batut, J., 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science 293, 668–672.

Gibson, K.E., Kobayashi, H., Walker, G.C., 2008. Molecular determinants of a symbiotic chronic infection. Annu. Rev. Genet. 42, 413–441.

Giuntini, E., Mengoni, A., De Filippo, C., Cavalieri, D., Aubin-Horth, N., Landry, C.R., Becker, A., Bazzicalupo, M., 2005. Large-scale genetic variation of the symbiosis-required megaplasmid pSymA revealed by comparative genomic analysis of Sinorhizobium meliloti natural strains. BMC Genomics 6, 158.

Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman, B.S., Cao, Y., Askenazi, M., Halling, C., Mullin, L., Houmiel, K., Gordon, J., Vaudin, M., Iartchouk, O., Epp, A., Liu, F., Wollam, C., Allinger, M., Doughty, D., Scott, C., Lappas, C., Markelz, B., Flanagan, C., Crowell, C., Gurson, J., Lomo, C., Sear, C., Strub, G., Cielo, C., Slater, S., 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science 294, 2323–2328.

Guo, X., Flores, M., Mavingui, P., Fuentes, S.I., Hernández, G., Dávila, G., Palacios, R., 2003. Natural genomic design in *Sinorhizobium meliloti*: novel genomic architectures. Genome Res. 13, 1810–1817.

Guzman, L.M., Belin, D., Carson, M.J., Beckwith, J., 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. 177, 4121–4130.

Hübner, P., Willison, J.C., Vignais, P.M., Bickle, T.A., 1991. Expression of regulatory *nif* genes in *Rhodobacter capsulatus*. J. Bacteriol. 173, 2993–2999.

Hanahan, D., 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166, 557–580.

Harrison, C.L., Crook, M.B., Peco, G., Long, S.R., Griffitts, J.S., 2011. Employing site-specific recombination for conditional genetic analysis in *Sinorhizobium meliloti*. Appl. Environ. Microbiol. 77, 3916–3922.

Hoess, R.H., Ziese, M., Sternberg, N., 1982. P1 site-specific recombination: nucleotide sequence of the recombining sites. Proc. Natl. Acad. Sci. U. S. A. 79, 3398–3402.

House, B.L., Mortimer, M.W., Kahn, M.L., 2004. New recombination methods for *Sinorhizobium meliloti* genetics. Appl. Environ. Microbiol. 70, 2806–2815.

Khan, S.R., Gaines, J., Roop, R.M., Farrand, S.K., 2008. Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. Appl. Environ. Microbiol. 74, 5053–5062.

Lambert, J.M., Bongers, R.S., Kleerebezem, M., 2007. Cre-*lox*-based system for multiple gene deletions and selectable-marker removal in *Lactobacillus plantarum*. Appl. Environ. Microbiol. 73, 1126–1135.

Langer, S.J., Ghafoori, A.P., Byrd, M., Leinwand, L., 2002. A genetic screen identifies novel non-compatible *loxP* sites. Nucleic Acids Res. 30, 3067–3077.

Lee, G., Saito, I., 1998. Role of nucleotide sequences of *loxP* spacer region in Cre-mediated recombination. Gene 216, 55–65.

Leigh, J.A., Reed, J.W., Hanks, J.F., Hirsch, A.M., Walker, G.C., 1987. *Rhizobium meliloti* mutants that fail to succinylate their calcofluor-binding exopolysaccharide are defective in nodule invasion. Cell 51, 579–587.

Liu, Z., Xie, Y., Zhang, X., Hu, X., Li, Y., Ding, X., Xia, L., Hu, S., 2016. Efficient construction of large genomic deletion in *Agrobacterium tumefaciens* by combination of Cre/*loxP* system and triple recombineering. Curr. Microbiol. 72, 465–472.

Milunovic, B., diCenzo, G.C., Morton, R.A., Finan, T.M., 2014. Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of *Sinorhizobium meliloti*. J. Bacteriol. 196, 811–824.

Missirlis, P.I., Smailus, D.E., Holt, R.A., 2006. A high-throughput screen identifying sequence and promiscuity characteristics of the *loxP* spacer region in Cre-mediated recombination. BMC Genomics 7, 73.

Murray, N.E., 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol. Mol. Biol. Rev. 64, 412–434.

Mus, F., Crook, M.B., Garcia, K., Garcia Costas, A., Geddes, B.A., Kouri, E.D., Paramasivan, P., Ryu, M.H., Oldroyd, G.E., Poole, P.S., Udvardi, M.K., Voigt, C.A., Ané, J.M., Peters, J.W., 2016. Symbiotic nitrogen fixation and challenges to extending it to non-legumes. Appl. Environ. Microbiol. 82, 3698–3710.

Niaudet, B., Jannière, L., Ehrlich, S.D., 1985. Integration of linear, heterologous DNA molecules into the *Bacillus subtilis* chromosome: mechanism and use in induction of predictable rearrangements. J. Bacteriol. 163, 111–120.

Oresnik, I.J., Liu, S.L., Yost, C.K., Hynes, M.F., 2000. Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. J. Bacteriol. 182, 3582–3586.

Peters, J.E., Thate, T.E., Craig, N.L., 2003. Definition of the *Escherichia coli* MC4100 genome by use of a DNA array. J. Bacteriol. 185, 2017–2021.

Reuber, T.L., Walker, G.C., 1993. Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. Cell 74, 269–280.

Sambrook, J., Russel, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., Pühler, A., 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145, 69–73.

Simon, R., Priefer, U., Pühler, A., 1983. A broad host range mobilization system for In vivo genetic engineering: transposon mutagenesis in gram negative bacteria. Nat. Biotechnol., 784–791.

Sternberg, N., Hamilton, D., 1981. Bacteriophage P1 site-specific recombination: I. Recombination between *loxP* sites. J. Mol. Biol. 150, 467–486.

Strand, T.A., Lale, R., Degnes, K.F., Lando, M., Valla, S., 2014. A new and improved host-independent plasmid system for RK2-based conjugal transfer. PLoS One 9, e90372.

Suzuki, N., Tsuge, Y., Inui, M., Yukawa, H., 2005. Cre/*loxP*-mediated deletion system for large genome rearrangements in *Corynebacterium glutamicum*. Appl. Microbiol. Biotechnol. 67, 225–233.

Vorhölter, F.J., Schneiker, S., Goesmann, A., Krause, L., Bekel, T., Kaiser, O., Linke, B., Patschkowski, T., Rückert, C., Schmid, J., Sidhu, V.K., Sieber, V., Tauch, A., Watt, S.A., Weisshaar, B., Becker, A., Niehaus, K., Pühler, A., 2008. The genome of *Xanthomonas campestris* pv. campestris B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. J. Biotechnol. 134, 33–45.

Wright, R., Stephens, C., Shapiro, L., 1997. The CcrM DNA methyltransferase is widespread in the alpha subdivision of proteobacteria, and its essential functions are conserved in *Rhizobium meliloti* and *Caulobacter crescentus*. J. Bacteriol. 179, 5869–5877.

Yu, B.J., Sung, B.H., Koob, M.D., Lee, C.H., Lee, J.H., Lee, W.S., Kim, M.S., Kim, S.C., 2002. Minimization of the *Escherichia coli* genome using a Tn5-targeted Cre/*loxP* excision system. Nat. Biotechnol. 20, 1018–1023.

Zhang, Y., Buchholz, F., Muyrers, J.P., Stewart, A.F., 1998. A new logic for DNA engineering using recombination in *Escherichia coli*. Nat. Genet. 20, 123–128.

6.4 A family of single copy *repABC*-type shuttle vectors stably maintained in the alpha-proteobacterium *Sinorhizobium meliloti*

Eigener Beitrag zur Studie:

I. Planung des pABC Vektor-Systems, der *in-vivo* Klonierungsstrategie sowie aller zugehörigen Experimente

II. Herstellung der pABC Modulvektor-Bibliothek

III. Stamm- und Plasmidkonstruktion, Assemblierung der pABCs in Zusammenarbeit mit MW

IV. Klonierung und Stammkonstruktion zur *in-vivo* Klonierung in Zusammenarbeit mit CH

V. Durchführung der Experimente in Zusammenarbeit mit MW und CH

VI. Mitverfasser des Manuskripts

(CH, Carina Happel; MW, Marcel Wagner)

A family of single copy *repABC*-type shuttle vectors stably maintained in the alpha-proteobacterium *Sinorhizobium meliloti*

Johannes Döhlemann[#], Marcel Wagner[#], Carina Happel, Patrick Sobetzko, and Anke Becker*

LOEWE Center for Synthetic Microbiology and Faculty of Biology, Philipps-Universität Marburg, Marburg, Germany

#These authors contributed equally.

*Correspondence: anke.becker@synmikro.uni-marburg.de

Abstract. A considerable share of bacterial species maintains segmented genomes. Plant symbiotic α-proteobacterial rhizobia contain up to six *repABC*-type replicons in addition to the primary chromosome. These low or unit-copy replicons, classified as secondary chromosomes, chromids or megaplasmids, are exclusively found in α-proteobacteria. Replication and faithful partitioning of these replicons to the daughter cells is mediated by the *repABC* region. The importance of α-rhizobial symbiotic nitrogen fixation for sustainable agriculture and *Agrobacterium*-mediated plant transformation as tool in plant sciences has increasingly moved biological engineering of these organisms into focus. Plasmids are ideal DNA-carrying vectors for these engineering efforts. Based on *repABC* regions collected from α-rhizobial secondary replicons, and origins of replication derived from traditional cloning vectors, we devised the versatile family of pABC shuttle vectors propagating in *Sinorhizobium meliloti* and *Escherichia coli*, and founded a modular plasmid library providing the elemental parts for pABC vector assembly. The standardized design of these vectors involves five basic modules: (1) *repABC* cassette, (2) plasmid-derived origin of replication, (3) RK2/RP4 mobilization site (optional), (4) antibiotic resistance gene, and (5) multiple cloning site flanked by transcription terminators. In *S. meliloti*, pABC vectors showed high propagation stability and unit-copy number. We demonstrated stable co-existence of three pABC vectors in addition to the two indigenous megaplasmids in *S. meliloti*, suggesting combinability of multiple compatible pABC plasmids. We further devised an *in vivo* cloning strategy involving Cre/*lox*-mediated translocation of large DNA fragments to an autonomously replicating *repABC*-based vector, followed by conjugation-mediated transfer either to compatible rhizobia or *E. coli*.

Keywords: plasmid cloning vehicle, Cre/*loxP*, *repABC, in-vivo* cloning, *Sinorhizobium*, *Agrobacterium*

Plasmid vectors are valuable tools in genetic engineering and synthetic biology. Most well established cloning vectors have specifically been selected and designed to be applied in *Escherichia coli*. Plasmids with a broader host range that can be transferred between different microorganisms and function in diverse genomic backgrounds represent a natural source of orthogonal DNA elements autonomously replicating in bacteria. The vegetative origin of replication (*oriV*) is the key element driving plasmid propagation by recruiting either host- or plasmid-encoded components for replication[1]. While medium to high copy plasmids are usually randomly distributed during cell division, low copy plasmids employ a segregation mechanism to assort plasmid copies to each daughter cell[2]. Plasmids can be considered as ideal DNA-carrying vectors with respect to the ease of plasmid engineering, modularization and standardization[3]. Yet, variations in copy number and propagation stability, as well as limitations in size and combinability are unfavorable properties of plasmid vectors, which motivated construction of secondary synthetic chromosomes[4–6] and genome engineering strategies relying on chromosomal integration of DNA constructs[7–9].

Several members of the Rhizobiales belonging to the class of α-proteobacteria are best known for their ability to establish a nitrogen-fixing symbiosis with legumes[10,11] or as plant pathogens of the genus *Agrobacterium* causing crown gall tumors on many vascular plants[12,13]. Advances in synthetic biology have increasingly moved utilization of genetic resources and processes underlying the microbe-plant interactions of these α-proteobacteria into focus. This includes the high value of symbiotic nitrogen fixation for sustainable agriculture[14], but also their biotechnological potential with respect to polysaccharide[15,16] and vitamine[17,18] production and *Agrobacterium*-mediated plant transformation[19].

In α-proteobacteria genome segmentation is a common phenomenon. Especially the fast-growing rhizobia harbor multipartite genomes which are usually composed of a main chromosome and additional low or unit-copy replicons classified as secondary chromosomes, chromids or megaplasmids[20,21], some of which even exceed 2 Mb in size. For instance, *Mesorhizobium loti* MAFF303099 possesses a chromosome and two megaplasmids (pMla, pMlb)[22], whereas *Sinorhizobium meliloti* Rm1021 harbors a chromosome, a chromid (pSymB), and a megaplasmid (pSymA)[23]. *Rhizobium etli* CFN42 even contains

six megaplasmids (p42a-f) beside its main chromosome[24], and *Agrobacterium tumefaciens* C58 maintains two chromosomes, a circular and a linear one, and two plasmids (pTi, pAt)[25,26]. The unrivaled variety of α-rhizobial genome architecture indicates a high flexibility in genome organization. Megaplasmids are faithfully inherited to the daughter cells. In *S. meliloti* Rm2011, duplication of the chromosomal origin of replication and secondary replicon *oriV*s is spatially and temporally uncoupled and occurs only once per cell cycle. Newly synthesized sister replicons partition in a strictly hierarchical order, starting with the chromosome and followed by pSymA and pSymB. This indicates a tight integration of these large secondary replicons into the cell cycle[27].

DNA replication origins of these extra-chromosomal replicons usually belong to the *repC* superfamiliy[21,28,29], named after the DNA replication initiator RepC, which appears to be unique to α-proteobacteria. In *repABC* family plasmids, *repC* including the cognate *oriV* forms an operon with *repA* and *repB*[29]. RepA and RepB belong to the ParAB family of partitioning proteins[29] mediating segregation of newly replicated *oriV*s. Segregation is initiated by specific binding of RepB to its cognate cis-acting centromere-like *repS* sites[29]. Highly reliable inheritance, even of dispensable *repABC*-type replicons such as most of the *R. etli* megaplasmids[30,31], can be attributed to proper function of the segregation machinery[32].

Apart from its role in DNA partitioning, RepA also functions as major transcriptional regulator of the *repABC* operon[21,28]. Stimulated by RepB, RepA binds to an operator upstream of its coding sequence resulting in strong transcriptional repression of *repABC*. The role of *repS* in autorepression of the *repABC* operon appears inconsistent and rather replicon-dependent. Whereas *repS* sites seem to be mandatory for effective autorepression of the *repABC* operon of pTi[33] derived from *A. tumefaciens*, *Rhizobium leguminosarum* derived RepA and RepB derivatives were sufficient for repression of P*repABC*$_{RtTA1}$ when heterologously expressed in *A. tumefaciens*[34]. Regulation of *repC* expression further occurs at post-transcriptional level via a small counter-transcribed RNA which is usually encoded by a locus between *repB* and *repC* (ctRNA)[28]. The ctRNA specifically targets the RepC mRNA, thereby covering its Shine-Dalgarno (SD) sequence, leading to strong translational inhibition[35,36]. Hence, replicon copy number is

supposed to be mainly transcriptionally controlled by RepAB-mediated autorepression of the *repABC* operon and post-transcriptionally by the ctRNA. However, further control mechanisms were identified for some *Rhizobium leguminosarum* and *A. tumefaciens repABC*-type plasmids whose copy number strongly increased in response to plant-released chemical signals and quorum-sensing transcription factors[37,38]. Strict regulation of *repABC* expression is the mechanistic basis of incompatibility groups allowing stable coexistence of several *repABC* replicons in the cell[21,29,36,39–41]. High variability of incompatibility factors, such as *repS* sites, the ctRNA and occasionally the C-terminal domain of RepC, could explain the compatibility of heterologous megaplasmids in related species[42].

Here, we explored the naturally occurring low or unit-copy number extrachromosomal elements of members of the *Rhizobiaceae* and *Phylobacteriaceae* to design the versatile family of modular standardized pABC plasmids, facilitating genetic engineering of this important group of bacteria. In *S. meliloti*, pABCs were shown to mimic megaplasmid-like characteristics, such as high propagation stability, unit-copy number and combinability of compatible RepABC-type plasmids, rendering them highly suitable for complex ectopic expression systems. We further devised an *in vivo* cloning strategy based on Cre/*lox*-mediated site-specific recombination and an IPTG-inducible engineered $oriV_{repABC}$, allowing for translocation of DNA fragments to an autonomously replicating *repABC*-based vector and conjugation-mediated transfer either to compatible rhizobia or to *E. coli*.

■ RESULTS AND DISCUSSION

Autonomous plasmid replication mediated by heterologous *repABC* regions in *S. meliloti*. Initially, we tested the ability of various *repABC* regions to mediate plasmid maintenance in *S. meliloti*. To this end, *repABC* regions derived from megaplasmids of *S. meliloti* (pSymA), *Rhizobium etli* (p42b, p42d, p42e, p42f), *Mesorhizobium loti* (pMla, pMlb), *Sinorhizobium medicae* (pSMED01, pSMED02), *Sinorhizobium fredii* (pNGR234b) and *Agrobacterium tumefaciens* (pTi) were inserted into the suicide vector pK18mob2 (Table 1), which is replicative in *E. coli*, but non-replicative in *S. meliloti*[43]. Plasmids carrying the *repABC* cassettes from *R. etli*, *M. loti* and *A. tumefaciens* were successfully

established in the *S. meliloti* Rm1021 wild type strain. In contrast, plasmids carrying the *repABC* regions derived from pSymA, pSMED01, pSMED02 and pNGR234b were not maintained in this strain, probably due to incompatibility between the indigenous megaplasmids and these heterologous replicons derived from closely related α-rhizobia. In line with this observation, the construct harboring the $repABC_{pSymA}$ cassette propagated in the pSymA-cured *S. meliloti* Rm1021 derivative Sma818, but not in the wild type. The identified set of eight heterologous *repABC* regions autonomously replicating in the *S. meliloti* Rm1021 wild type strain provided the foundation for construction of the family of pABC vectors.

pABC vector design. Simple design concepts and modularity facilitating assembly and exchange of functional modules have become standard in vector development for applications in synthetic biology.

Figure 1 depicts the blueprint of the pABC vector family composed of five basic modules: oriVSm (*repABC* cassette for propagation in *S. meliloti*), oriVEc (plasmid-derived origin of replication for propagation in *E. coli*), oriT (RK2/RP4 mobilization site), AR (antibiotic resistance gene cassette), and synTer-MCS (multiple cloning site flanked by transcription terminators).

One of our motivations for developing this vector family is the possibility of co-existence of multiple pABC replicons derived from compatible megaplasmids in one cell. Natural occurrence of α-rhizobial species with up to six compatible *repABC*-type replicons indicate the high potential of combinability.

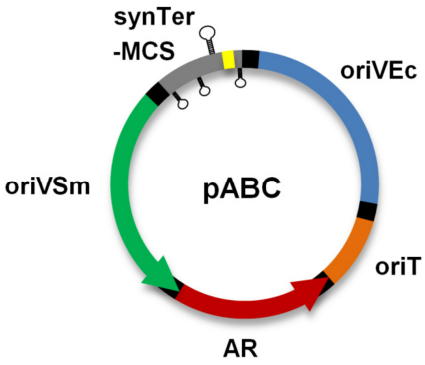


Figure 1. pABC vector composed of five module library parts. Modules oriVSm and oriVEc mediate propagation in *S. meliloti* and *E. coli*, respectively. To ensure proper transcriptional control of the *repABC* operon, the multiple cloning site (yellow box) of the synTer-MCS module, serving as insertion site for the gene load, is flanked by transcriptional terminators (loop structures). A common RK2/RP4 mobili-zation site (module oriT) enables con-jugal transfer of pABCs from *E. coli* to *S. meliloti*. For reliable selection in both organisms standardized antibiotic resis-tance cassettes were developed (module AR). Each module is flanked by linker sequences for standardized LCR-based assembly (black boxes). In the pABC design, order and orientation of modules is intended to minimize mutual influ-ences of adjacent regions.

Previously, for high-throughput plasmid integration study[44], we determined 220 bp sharing at least 80% sequence similarity with the target region as the minimal size promoting homologous recombination at a significant frequency in *S. meliloti* Rm1021. To limit the possibility of homologous recombination between co-residing members of the pABC vector family, parts were designed on the premise of avoiding high and extended sequence identity (Table S1). Multiple parts per module allow for assembly of combinable pABC vectors. The only exception is the oriT module that comprises parts with extended sequence identity. It is therefore not recommendable to combine pABC plasmids with this module in one cell.

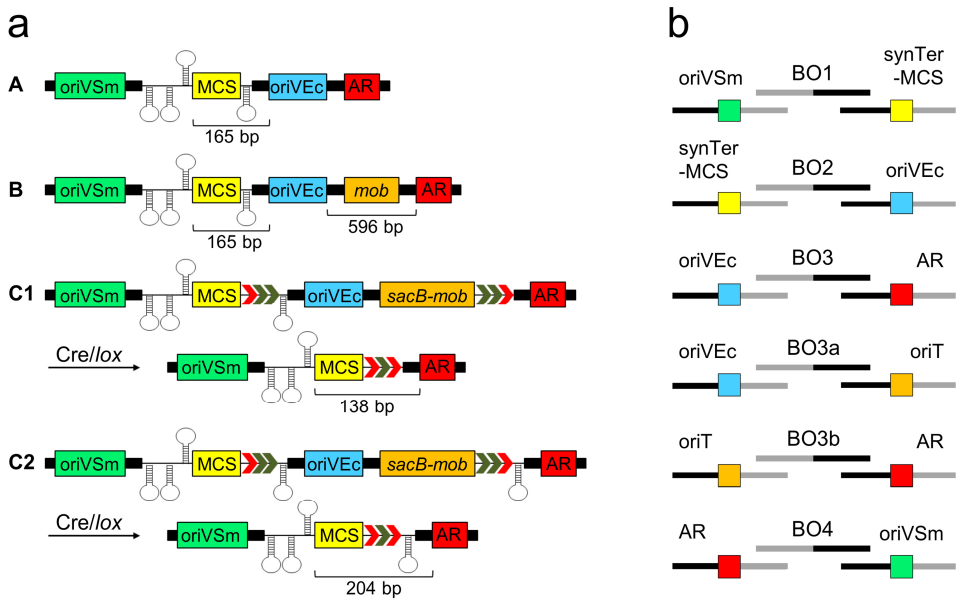


Figure 2. Concept of the pABC assembly system. (a) Detailed scheme of pABC design types A to C. The size of design-derived homologous regions potentially occurring in multi-pABC systems is denoted. Setups C1 and C2 differ in the number of transcription terminator sequences (loop structures) protecting the MCS after Cre-mediated deletion of modules oriVEc and oriT. (b) Module-specific linker sequences (grey and black lines) facilitate LCR-based assembly of module parts with standardized bridging oligonucleotides BO1 to BO4. Integration of the oriT module requires BO3a and BO3b.

Three basic pABC types were devised (Figure 2a). pABCs of type A, lacking the oriT module, have to be introduced to *S. meliloti* by electroporation. pABCs of type B contain an oriT module (pABC-mob) enabling conjugal transfer. Hence, unintended co-integration by homologous recombination between type B pABC vectors through the ~500 bp *mob* site appears possible. This is circumvented by design type C

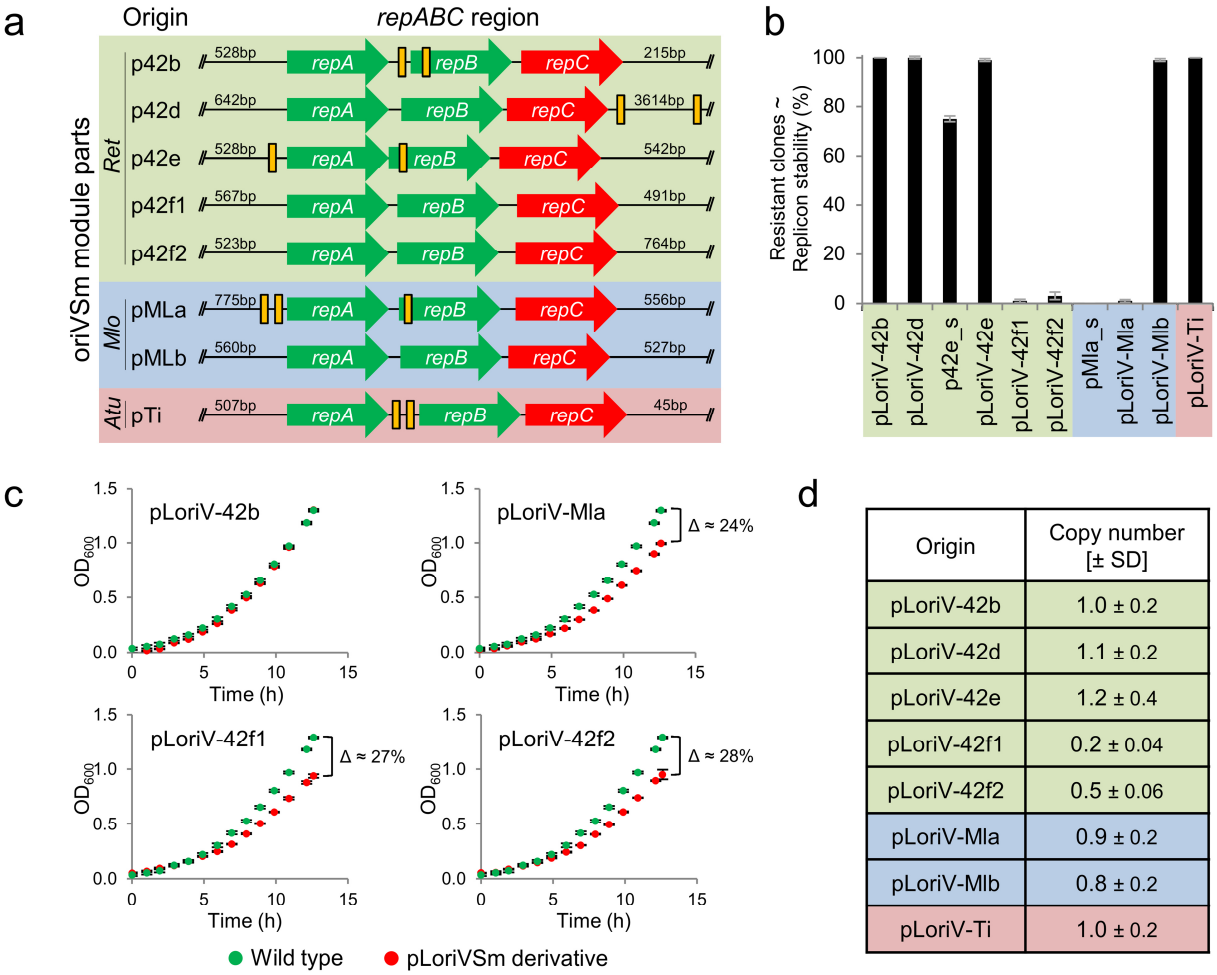
which offers the possibility of excision of the oriVEc and oriT modules by Cre/*loxP*-mediated site-specific recombination in *S. meliloti*[45,46].

LCR (ligase cycling reaction)[47] is our procedure of choice for restriction enzyme-independent assembly of the building blocks to pABC vectors. LCR was shown to efficiently assemble up to ten 2 kb DNA fragments[47]. Library plasmids serving as template for PCR amplification of the 23 parts generated in our study were constructed (Table 1). Module-specific linker flank each building block to enable LCR-based assembly promoted by standardized bridging oligonucleotides (Figure 2b, Table S2). Following *in-vitro* assembly, *E. coli* DH5α was directly transformed with the purified LCR reaction mix and correct assembly of pABC vectors was verified by DNA sequencing. Usually more than 75% of tested clones exhibited correct assembly, indicating a suitably efficient assembly system.

Library parts for each module (Tables 1 and S3) and pABC vectors (Tables 2 and S4) generated in this study are described in the SEVA-SIB database (http://seva.cnb.csic.es)[48] (pending). In order to translate individual pABCs into the SEVA-SIB collection, a unique identifier was assigned to each insert in accordance with the standard SEVA nomenclature (Tables S3 and S4).

Parts characterization. oriVSm module: The eponymous element of the pABC vector family is the *repABC* region which naturally acts as replication origin of megaplasmids of various α-proteobacteria[21,28]. *repABC* cassettes of different megaplasmids were integrated into pK18mob2 derivative pLoriVSm resulting in library module vectors which provide the *repABC* cassettes of *R. etli* p42b,d,e,f (pLoriV-42b, -42d, -42e, -42f1, -42f2), *M. loti* pMla,b (pLoriV-Mla, -Mlb) and *A. tumefaciens* pTi (pLoriV-Ti).

Figure 3a depicts the genetic structure of the eight *repABC* regions, which constitute the basis of the oriVSm module parts used in this study. To test for autonomous propagation and stability in absence of selective pressure, *S. meliloti* Rm1021 was transformed with the pLoriVSm library plasmids. Restriction enzyme-digested pLoriVSm plasmids purified from the *S. meliloti* transformants showed the expected electrophoresis banding patterns (Figure S1), verifying autonomous propagation of these replicons in *S. meliloti*.



Origin	Copy number [± SD]
pLoriV-42b	1.0 ± 0.2
pLoriV-42d	1.1 ± 0.2
pLoriV-42e	1.2 ± 0.4
pLoriV-42f1	0.2 ± 0.04
pLoriV-42f2	0.5 ± 0.06
pLoriV-Mla	0.9 ± 0.2
pLoriV-Mlb	0.8 ± 0.2
pLoriV-Ti	1.0 ± 0.2

Figure 3. Characterization of oriVSm module parts. (a) *repABC* cassettes analyzed in this study. Regions are derived from *R. etli* (*Ret*), *M. loti* (*Mlo*) and *A. tumefaciens* (*Atu*) plasmids. Since functional elements, such as partitioning sites (*repS*, yellow boxes) or the promoter of the antisense RNA incα potentially are located outside of the *repABC* operon, *repABC* cassettes were expanded several hundred base pairs up- and downstream of the coding regions. Although a suicide vector carrying the *repABC* operon of p42d plus 500 bp downstream of the *repC* stop codon was able to propagate in *R. etli* CFNX101[74], the downstream region was generously expanded in order to include a further predicted *repS* site[75]. (b) Stability assay of pLoriVSm derivatives. *S. meliloti* Rm1021 carrying pLoriV-42b, -42d, 42e, -Mla, -Mlb, or -Ti, or the shorter constructs p42e_s or pMla_s were grown for 72 h in non-selective TY medium and diluted every 12 h to an OD_{600} of ~0.05. Single colonies were examined for antibiotic resistance (n= 100 for pLoriVSm derivatives; n= 50 for p42e_s and pMla_s). Error bars indicate the standard deviation calculated from three replicates. (c) Growth of *S. meliloti* Rm1021 carrying selected pLoriVSm plasmids compared to MWSm8 (Wild type, Km^r). Cultures were adjusted to an OD_{600} of 0.1 and grown for 12 h in selective TY medium. Error bars indicate the standard deviation calculated from four replicates. (d) qPCR-based copy number determination of pLoriVSm derivatives (-42b, -42d, -42e, -42f1, -42f2, -Mla, -Mlb and -Ti) in exponentially growing *S. meliloti*. Measurements are based on four replicates. SD: Standard deviation.

pLoriV-42b, -42d, -42e, -Mlb and -Ti exhibited highly stable propagation despite cultures were kept in exponential growth for 72 h in non-selective rich medium (Figure 3b). Expanding the 3'UTR of the *repABC*$_{p42e}$ region from 26 bp (p42e_s) to 542 bp (pLoriV-42e) resulted in stable propagation, indicating that genetic elements required for faithful replication or segregation of these replicons can be located further downstream of the *repABC* operon. In contrast, plasmids harboring *oriV*$_{p42f1}$, *oriV*$_{p42f2}$ and *oriV*$_{pMla}$ derived cassettes were virtually lost in the same period of time. Expanding the 3'UTR of *repABC*$_{pMla}$ from 45 bp (pMla_s) to 556 bp (pLoriV-Mla) did not improve plasmid maintenance. Slight incompatibility between *repABC*$_{pMla}$ and the *oriV* regions of the indigenous megaplasmids may account for this viable but instable coexistence. Cross-interactions between components of indigenous and heterologous RepABC systems may impair regulatory control of the *repABC* operons or function of the gene products in replication initiation or plasmid segregation. This may affect cell cycle progression and diminish cell growth. Indeed, pLoriV-42f1, -42f2 and -Mla negatively affected *S. meliloti* cell growth, whereas stably propagating plasmids bearing *oriV* cassettes from p42b, p42e, pMlb and pTi had no significant influence on growth (Figure 3c and S2). This is supposed to be a critical prerequisite for a multi-pABC setup under non-selective conditions since a negative effect on cell growth would promote plasmid loss.

We aimed at preserving the unit- or low copy number of megaplasmids in the free-living state as important attribute of pABC vectors. The copy number of pLoriVSm plasmids in *S. meliloti* was determined by qPCR and data were normalized to pK18mob2-*repABC*$_{pSymA}$, which in a calibration experiment was shown to exhibit the same copy number as the chromosome (Table S5). The stably propagating pLoriV library plasmids roughly matched a unit-copy number, whereas instable plasmids showed values below 1 (Figure 3d and Table S6). The latter can be ascribed to the applied method. Since qPCR based copy number determination was performed on crude cell extracts, a significant number of pLoriV plasmid-negative cells could take account for the low values determined for plasmids pLoriV-42f1 and -42f2 (Figure 3d).

Considering stability and copy number, the oriVSm module vectors pLoriV-42b, -42d, -42e, -Mlb and -Ti exhibited the desired megaplasmid-like characteristics and thus appear to be highly suitable for construction of the pABC vector system.

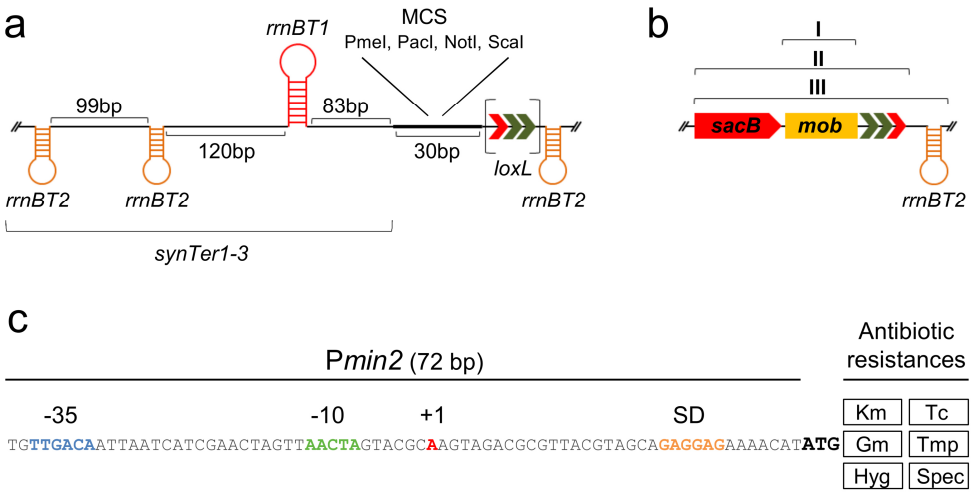


Figure 4. Setup of pABC library modules. (a) The synTer-MCS module plasmids provide three ~550 bp *synTer* regions composed of Rho-independent *E. coli* terminators *rrnBT1* (90 bp) and *rrnBT2* (60 bp) separated by unique synthetic sequences. Terminator sites (loop structures) are intended to protect the oriVSm module and the MCS, which provides unique recognition sites for rare cutters. pLsynTerX-lox plasmids are further equipped with a left arm mutated *loxL* site (arrows in brackets). (b) The oriT module comprises a 460 bp *mob* site (I). For Cre/*lox* applications the region is extended by *sacB-loxR* (II) and a further *rrnBT2* transcription terminator (III). (c) Module AR contains the short P*min2* fragment conferring resistance to kanamycin (Km), tetracycline (Tc), gentamicin (Gm), trimethoprim (Tmp), hygromycin (Hyg) or spectinomycin (Spec) in *E. coli* and *S. meliloti*.

synTer-MCS module: This module provides a multiple cloning site offering recognition sites for the rare-cutting restriction enzymes PmeI, PacI, and NotI. Transcriptional read-through into the *repABC* operon was found to interfere with its strict native expression control. The MCS module was therefore additionally equipped with flanking transcription terminators to protect the oriVSm module and the MCS position from transcriptional read-through emanating from sequences inserted into the MCS or adjacent modules. To this end, termination regions flanking the MCS were designed based on the *E. coli rrnBT1* and *rrnBT2* terminators (Figure 4a) and cloned into pK18mob2 to generate the pLsynTer library plasmids (Table 1). Effective transcription termination was verified by interrupting expression of an *eGfp* reporter

gene from a strong constitutive promoter by inserting the *synTer*1 to 3 cassettes downstream of the promoter (Figure S3).

oriVEc and oriT modules: The oriVEc module enables *E. coli* based cloning of pABCs. However, construction of standardized library plasmids carrying a further *E. coli* replication origin (besides the native $oriV_{pMB1}$) failed. For this reason, oriVEc parts harboring $oriV_{pMB1}$, $oriV_{p15A}$ and $oriV_{pSC101*}$ were PCR-amplified from non-library plasmids by use of extended primers carrying the module specific linker sequences. The oriT module provides an RP4-derived *mob* site which mediates conjugal transfer from *E. coli* to a wide range of Gram-negative bacteria[49] including *S. meliloti*. This module is available on the pLoriT library plasmids (Table 1).

Low variability of oriVEc and oriT module parts is associated with a reduced combinability of pABC vectors since homologous regions provoke unintended replicon fusions. For this reason we equipped the synTer-MCS and oriT modules with Cre recognition (*lox*) sites, resulting in the pLsynTer-x-lox and pLoriT-x-lox library plasmids, respectively (Figure 4a-b, Table 1). These *lox* sites allow Cre-mediated deletion of these modules from the pABC vectors in the α-rhizobial host strain SmCreΔhsdR as demonstrated in Figure S4.

AR module: This module comprises selection markers conferring resistance to the antibiotics gentamicin, hygromycin, kanamycin, tetracycline, trimethoprim, and spectinomycin in *E. coli* and *S. meliloti* (Figure S5). For the AR module, we chose a standardized design with the 72 bp P*min2* promoter driving expression of the resistance genes (Figure 4c). This promoter was selected from P*aacC1* and P*trp* derivatives (Figure S6). AR module parts are available on the pLAR library plasmids (Table 1).

Characterization of pABC shuttle vectors. We constructed 14 pABCs representing design types A to C utilizing all oriVSm module parts described above (Tables 2 and S7). pABCa-c are based on constructs lacking the standard linker sequences flanking the module library parts. They basically match the structure of pABC1 to 6 which are composed of library parts assembled via module-specific linker sequences and include an additional transcription terminator downstream of the MCS (Figure S7a). Dependent on the pABC design, transformation of *S. meliloti* Rm1021 was carried out via electroporation

or *E. coli* S17-1 mediated conjugal transfer. Following electroporation, transformation rates of *S. meliloti* with pABCs were massively higher than with the common replicative vectors pSRKGm and pPHU231[45]. Compared to pSRKGm, pABCa, pABCb and pABCc achieved ~7600-, 93-, and 23-fold higher transformation efficiencies, respectively, making co-transformation of *S. meliloti* with up to three pABCs possible (Figure S7b). Successful introduction of pABCs into *S. meliloti* and autonomous propagation was verified by plasmid purification from transconjugants and restriction enzyme digest-derived fragment patterns (Figure S8). Replicon fusion or genomic integration of pABCs into the genome were not observed.

Inheritance of pABC vectors was studied by flow cytometry-based single cell analysis applying *S. meliloti* strains carrying pABC-*eGFP* derivatives (Table S8). In contrast to the preliminary selection marker-based assay of maintenance of the pLoriVSm library plasmids (Figure 3b), which tends to preselect fast growing cells, the flow cytometry-based assay allows for unbiased determination of the ratio of pABC-positive and -negative cells and is highly sensitive (Figure S9). Cultures were incubated in non-selective rich medium and diluted twice a day to allow for continuous growth in the range of OD_{600} of 0.05 to 1.0.

Consistent with the properties of the corresponding pLoriVSm plasmids, derivatives of pABCa ($oriV_{pMlb}$), pABCb ($oriV_{p42b}$), pABCc-mob ($oriV_{p42d}$), pABC1 ($oriV_{p42b}$) and pABC4 ($oriV_{p42e}$) were remarkably stably propagated over the time course of 96 h (Figure 5a i). After 24 h even ~99% of measured cells were eGFP-positive and therefore supposed to maintain the respective pABC replicon. After 96 h of incubation still 97% of cells were positive, indicating a highly reliable propagation of the pABC derivatives in *S. meliloti* Rm1021. In contrast, *eGFP* expressing derivatives of pABC2 ($oriV_{pMla}$) and the low-copy plasmid pPHU231 showed a pronounced plasmid loss of ~ 20% after 24 h (Figure 5a ii). After 96 h, the proportion of positive cells even decreased to 13% and 39%, respectively. Notably, flow cytometry revealed dramatic variations in propagation stability between biological replicates of clones bearing the $oriV_{pTi}$-based pABC3-*eGFP* (99%, 49% and 2% positive cells after 48 h) (Figure 5a ii).

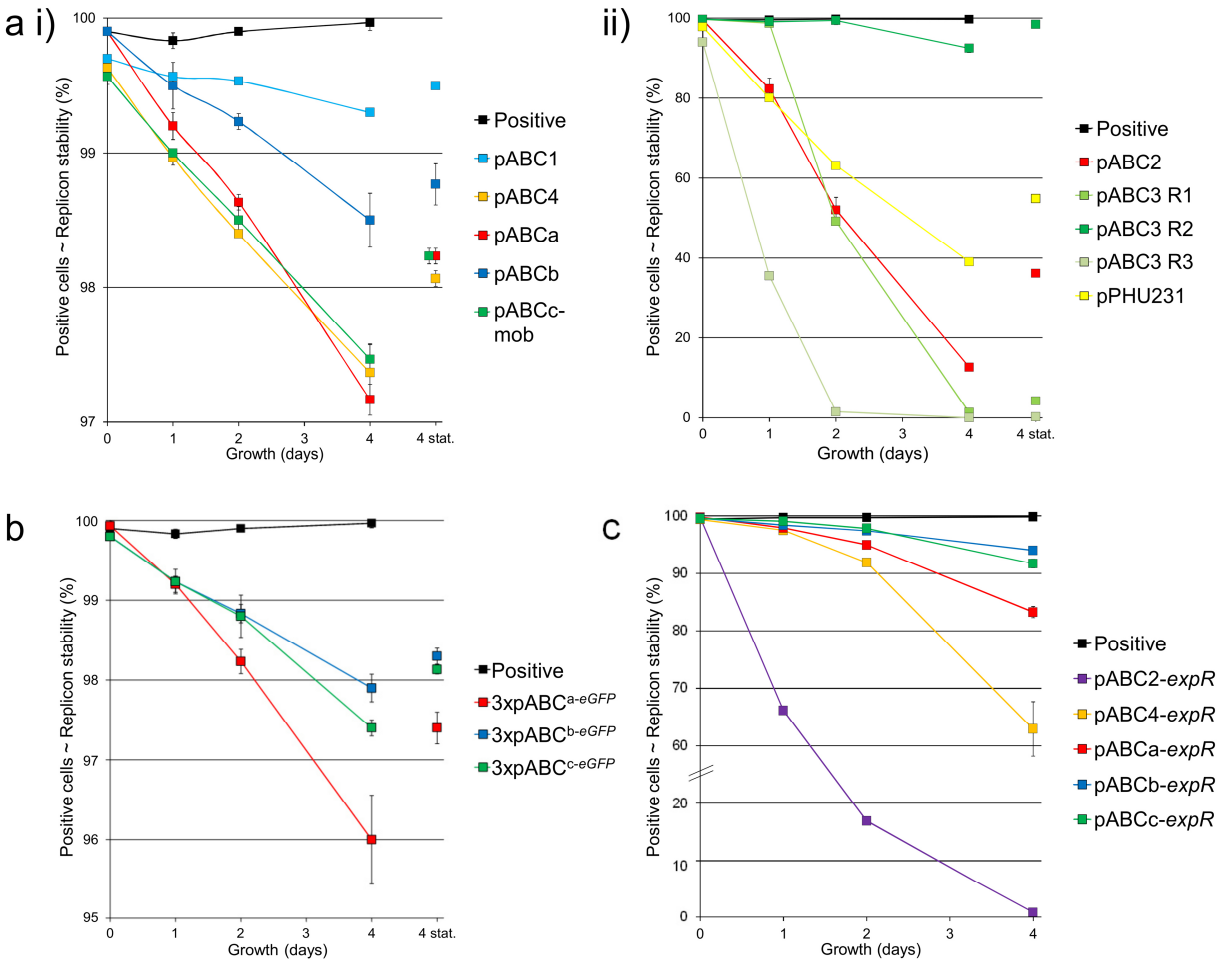


Figure 5. Inheritance stability of pABCs in *S. meliloti* Rm1021. *S. meliloti* carrying pABC-*eGFP* (-*expR*) derivatives were grown in non-selective TY medium at 30°C for 96 h. During the time course, cultures were diluted every 12 h to an OD_{600} of ~ 0.05. eGFP signal of 10.000 cells was measured by flow cytometry at time point 0 and after 1, 2, and 4 days. Additionally, after initial 24 h subcultures were incubated for 48 h until stationary phase was reached, and re-inoculated at day 3 (4 stat.). Positive: *S. meliloti* MWSm37 (genomically integrated *eGFP*). Error bars indicate standard deviation calculated from triplicates. (a) Maintenance of single pABC-*eGFP* derivatives. i) Stable pABC-*eGFP* constructs which did not deceed 97 % propagation stability. ii) Replicon stability of instable pABC-*eGFP* derivatives and low copy plasmid pPHU231-*eGFP*. Due to heterogeneous behavior of pABC3-*eGFP* individual replicates are shown (R1-3). (b) Replicon stability of pABCs in a triple pABC system. 3xpABC$^{a\text{-}eGFP}$: pABCa-*eGFP*, pABCb, pABCc-mob; 3xpABC$^{b\text{-}eGFP}$: pABCa, pABCb-*eGFP*, pABCc-mob; 3xpABC$^{c\text{-}eGFP}$: pABCa, pABCb, pABCc-mob-*eGFP*. (c) *S. meliloti* carrying pABC-*eGFP*-*expR* derivatives.

Mutation of the cognate *repABC* cassette was excluded by re-sequencing. However, we can't exclude mutations in other pABC modules or the *S. meliloti* background genome. Previously, regulatory

mechanisms affecting replication and thereby the copy number of pTi in *A. tumefaciens* have been reported[50]. The heterogeneity in propagation stability observed in our study and variations in copy number of pTi characterize $oriV_{pTi}$ as unsuitable part for construction of stably propagating single copy pABC vectors in *S. meliloti*. The pABC maintenance assay identified $oriV_{pMlb}$, $oriV_{p42b}$, $oriV_{p42d}$, and $oriV_{p42e}$ as highly suitable parts for generation of *S. meliloti* pABC vectors.

In order to rule out mutual influences of the *repABC* cassettes in multi-pABC setups we further studied replicon stability in *S. meliloti* strains simultaneously carrying two or three pABCs. pABCa to c retained highly stable propagation (>96%) throughout the time course and achieved stability values which were only marginally lower compared to the strains carrying only one of these pABCs (Figure 5b). pABC3 and 4 were combined to study the influence of an instably propagating pABC on maintenance of a stably inherited pABC. In combination, pABC3 and pABC4 showed the same properties as individually, namely biological replicates show great variation in maintenance of pABC3 while pABC4 was stably inherited (Figures S10). In accordance with the previous growth studies of strains harboring stably propagating pLoriVSm plasmids (Figure 3c and S2), the corresponding pABCs did not affect *S. meliloti* cell growth even when present in combination (Figure S11).

We further tested maintenance of stably propagating pABCs under more challenging conditions by loading of pABC vectors with the *expR* gene, whose presence restrains growth of *S. meliloti*[51]. In the reference strain Rm1021, *expR* is disrupted by an insertion element[52]. Complementation of this mutation results in a strong increase in galactoglucan production since the LuxR-type regulator ExpR is a key activator of biosynthesis of this exopolysaccharide [52–56]. *expR* under the control of P*lacZ* was integrated into pABCs based on *repABC* cassettes derived from pMlb (pABCa-*eGPF-expR*), p42b (pABCb-*eGPF-expR*), p42d (pABCc-*eGPF-expR*), p42e (pABC4-*eGPF-expR*) and pMla (pABC2-*eGPF-expR*). Constitutive expression of *expR* significantly decreased cell growth and resulted in a mucoid colony phenotype, indicating a proper function of the *expR* cassette in the pABCs (Figure S12a). We compared maintenance of *expR*-positive and *expR*-negative pABCs reliably and unreliably propagating in *S. meliloti*. After 96 h, the *expR*-positive derivative of the unreliably propagating pABC2 (pABC2-*eGFP-*

expR) was virtually lost (Figure 5c), whereas the derivative lacking *expR* (pABC2-*eGFP*) was still present in 17% of the cells (Figure 5a ii). Maintenance of pABCs characterized as very reliably inherited to the daughter cells was less affected by the presence of *expR*. After 96 h, 63%, 83%, and 92 to 94% of cells contained pABC4, pABCa, and pABCb,c, respectively. Nevertheless, derivatives of pABCa to c and pABC4 were reliably maintained with 98 to 99% positive cells after 24 h, despite of carrying a functional copy of *expR* throughout the entire time course (Figure S12b). Hence, vectors bearing *oriV* cassettes derived from p42b, p42d, p42e and pMlb exhibit remarkable inheritance even when imposing a fitness burden on the cell.

To validate the unit copy number of stably propagating pABCs and to rule out potential morphological aberrations of pABC-bearing cells, cells carrying fluorescently labeled pABC derivatives were microscopically analyzed. To this end, *Yersinia pestis* and fluorescent repressor operator system (FROS) derived *parS*/mChr-ParB[57] and *tetO*/TetR-YFP[58] pairs were applied as taurine-inducible dual color label system. Localization patterns of *parS*/mChr-ParB and *tetO*/TetR-YFP clusters implied single copy replication of pABCa to c and pABC1, 2 and 4, since the number of fluorescent foci never exceeded two per cell (Figure S13a). This observation is in accordance with the copy number of pLoriVSm constructs determined by qPCR (Figure 3d). Cell populations of the *S. meliloti* strains with the unreliably inherited pABC5 ($oriV_{p24f1}$) and pABC6 ($oriV_{p42f2}$) vectors contained a substantial proportion of cells lacking fluorescent foci, indicating loss of the respective vector (Figure S13b). Aberrant cell morphologies, such as branching and elongation, were not observed. Previously, we observed that the *repA2B2C2* region derived from pSymA is sufficient to confer the spatiotemporal pattern of this megaplasmid to a small plasmid[27]. Although, spatiotemporal dynamics and integration of pABC into the cell cycle must be examined in more detail in future studies, microscopic analyses suggest a megaplasmid-like coordination with subpolar localization in small cells at the beginning of the cell cycle and in elongated pre-divisional cells (Figure 6a).

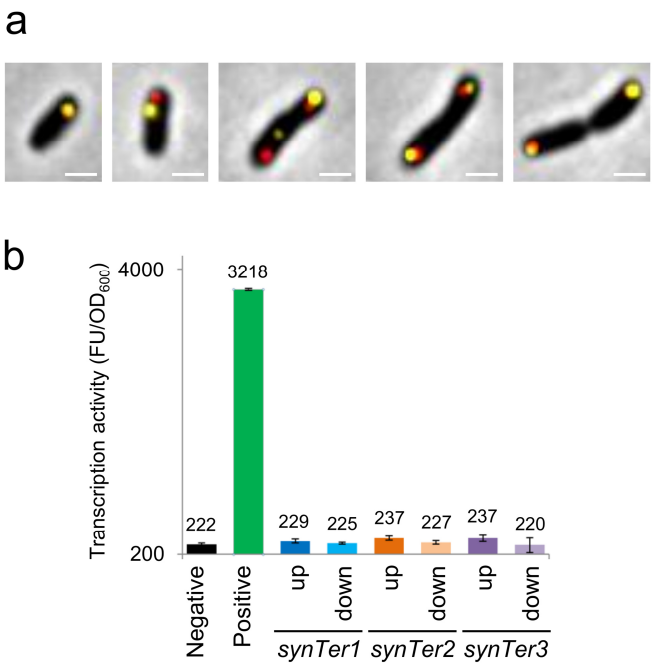


Figure 6. Examination of pABCs in *S. meliloti*. (a) Fluorescence microscopy of *S. meliloti* JDSm106 carrying pABCa-*tetO,* pABCb and pABCc-*parS*$_{Yp}$. Expression of reporters *tetR-YFP* (yellow) and *mChr-parB*$_{Yp}$ (red) was induced with 15 mM taurine for 4 h. Snapshots show cells at different stages of the cell cycle. Scale bar: 1 µm. (b) eGFP fluorescence of *S. meliloti* Rm1021 carrying pABC1-*eGFPf/r* (*synTer1*), pABC1-ST2-*eGFPf/r* (*synTer2*) and pABC1-ST3-*eGFPf/r* (*synTer3*) was taken as measure for the transcription activity in the MCS of library-derived pABCs. Promoter-less *eGFPf/r* cassettes were either forwardly (*eGFPf*) or reversely (*eGFPr*) (corresponding to up/down orientation) integrated into the MCS. Negative: *S. meliloti*/pABC1; Positive: *S. meliloti*/pABC1-*eGFP*. Error bars indicate standard deviation calculated from three replicates. FU: fluorescence units.

The main criteria for the pABC vectors are considered to be high propagation stability, no effect on cell growth and a single copy number on the one hand, and a well-defined insulated cloning site on the other hand. Insulation was tested by measuring fluorescence derived from promoter-less reporter gene insertions into the cloning site (Figure 6b and S14). Thus, pABCa, pABCb, pABCc, pABC1 and pABC4 are supposed to be the most reliable candidates for applications in *S. meliloti*.

***repABC* mediated *in-vivo* cloning in *S. meliloti* (ABC cloning).** Several methods have been developed which enable direct cloning of large gene clusters, including transformation-associated recombination (TAR) in *Saccharomyces cerevisae*[59–62], RecET-mediated linear-plus-linear homologous recombination (LLHR) in *E. coli*[63], and φBT1 integrase-mediated site-specific recombination in

Streptomyces species[64]. The Cre/*loxP* plus BAC strategy[65] and *oriT*-directed cloning[66] are well established *in-vivo* approaches which facilitate cloning of large DNA fragments derived from α-rhizobia. The former strategy is based on Cre/*loxP*-mediated circularization of the DNA fragment of interest bearing a BAC *oriV*, which is purified from the crude cell culture and introduced to *E. coli* by electroporation[65]. The latter strategy involves Flp-mediated excision of FRT site bordered DNA fragments including an *oriT* site and an *E. coli* plasmid *oriV*, instantly followed by mobilization of the excised DNA to *E. coli*[66]. These direct cloning methods have proven to be very useful. However, they do not allow reliable maintenance of excised DNA in propagating donor cells, but require either DNA purification from treated cells and subsequent transformation of the crude mix into a suitable host or excision and transfer in a single step and therefore entail reduced efficiency and increased screening efforts. Properties of α-rhizobial megaplasmids such as single copy number, faithful inheritance to daughter cells and large size make *repABC*-type *oriV*s ideal tools for *in-vivo* cloning strategies. We devised an *in-vivo* cloning strategy involving Cre/*lox* mediated excision of the target region, followed by stable extrachromosomal establishment of the excised region on an autonomously replicating *repABC*-type replicon, and demonstrated its feasibility by translocation of the ~24 kb *exo* gene cluster from pSymB to a *repABC*-type replicon (Figure 7).

This strategy requires genomic integration of a *repABC*-type *oriV* region adjacent to the region of interest. To enable stable integration, we constructed a LacI-silenced and IPTG-inducible $repABC_{pMlb}$ operon by insertion of a *lac* operator sequence into the *repABC* promoter region (Figure 7a). Constitutive *lacI* expression strongly affected the capacity of pMlb_*lacO* bearing $repABC_{pMlb}$-*lacO* to transform *S. meliloti*, most probably due to repression of the *repABC* operon (Figure S15a). Whereas establishment of pMlb_*lacO* in *S. meliloti* carrying the medium copy plasmid pSRKGm as LacI donor was completely abolished, *S. meliloti* harboring the single copy plasmid pABCb-*lacI* exhibited still sufficiently reduced transformation rates.

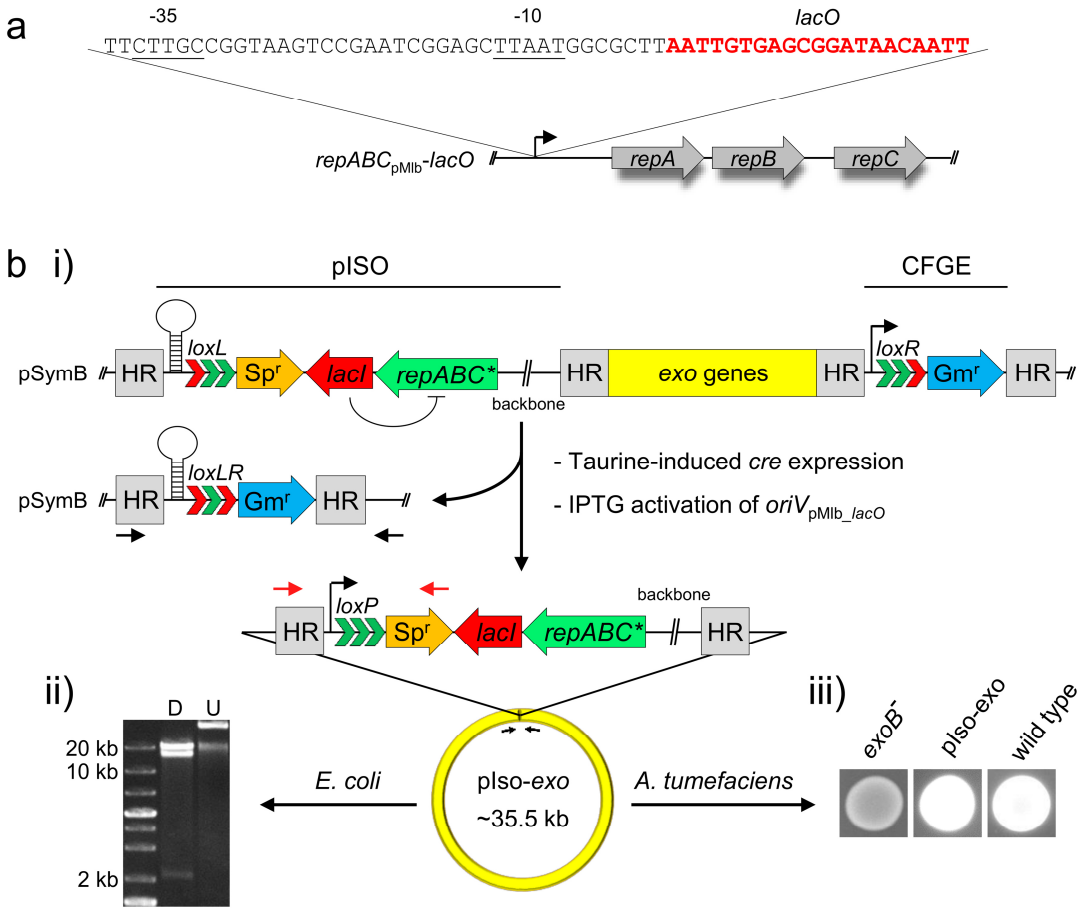


Figure 7. *repABC* mediated *in-vivo* cloning in *S. meliloti* (ABC cloning). (a) The promoter region of *repABC*$_{pMLb}$ was predicted according to the consensus motifs of *S. meliloti* σ^{70} promoters (-35 and -10 elements are underlined). A single *lacO* box (red letters) was integrated at the predicted transcription start site (+1). (b) ABC cloning of the *exo* gene cluster in *S. meliloti* SmCre. i) A gentamicin resistance cassette carrying a right arm-mutated *loxR* site (*Pmin2-loxR-aacC1*) was integrated downstream of the gene cluster via cloning-free genome editing[45] (CFGE). Subsequently, pISO bearing *repABC*$_{pMlb}$-*lacO* (*repABC**) was integrated via a ~550 bp homologous region (HR) upstream of the gene cluster, resulting in *S. meliloti* SmCre_exo-IN. Taurine induction of Cre expression and simultaneous IPTG activation of *oriV*$_{pMLb\_lacO}$ gave rise to SmCre_exo-OUT which carries the relocated ~35.5 kb region (pIso-exo) comprising the entire *exo* gene cluster. Both the deletion site on pSymB and the fusion site on pIso-exo were PCR-amplified with primers 94+Rev and 680+128 (black arrows), respectively, and sequencing of PCR products confirmed proper Cre-mediated recombination. Three arrows indicate *lox* sites with native (green) or mutated arms (red). ii) pISO-*exo* was transferred to *E. coli* via triparental mating, resulting in *E. coli*/pIso-exo. Purifed plasmid DNA was digested with NheI, resulting in fragments of 2.03 kb, 13.8 kb and 19.7 kb fragments (D). Sequencing of the fusion site comprising *loxP* and *aadA1* with primers 680 and 128 (red arrows) further confirmed successful cloning. U: Undigested plasmid DNA. iii) Calcofluor fluorescence assay. pISO-exo was transferred to the exopolysaccharide-deficient *A. tumefaciens* C58 exoB mutant (*exoB*$^-$) via triparental mating, giving rise to *A. tumefaciens* exoB/pIso-exo complemented for exopolysaccharid production (pIso-exo), and thus showing UV-induced fluorescence on Calcofluor-containing medium. Wild type: *A. tumefaciens* C58.

Thus, *lacI* expression from a single copy locus achieved sufficient inactivation of *repABC*$_{pMlb}$-*lacO*. In agreement with these findings, re-growth of the *S. meliloti* transconjugants on IPTG-supplemented medium restored high transformation rates, indicating an IPTG-dependent activation of the engineered *oriV*.

Figure 7b (i) depicts the genomic situation of *S. meliloti* SmCre_exo-IN bearing pIso and P*min2-loxR-aacC1* integrated up- and downstream of the *exo* gene cluster, respectively. Whereas the downstream fragment constitutes a functional gentamicin resistance cassette which is separated from the *exo* gene cluster by the Cre recognition site *loxR*, pIso provides a *loxL* site which is preceded by a terminator sequence preventing transcription of the promoter-less spectinomycin resistance gene *aadA1*. pIso further offers a constitutively expressed *lacI* gene encoding the transcriptional repressor which mutes the adjacent *repABC*$_{pMlb\_lacO}$ operon in absence of IPTG. This repression of *oriV* activity enables integration of pIso into the genome. Accordingly, the efficiency of *S. meliloti* transformation with pIso was ~3000-fold decreased when mating mixtures were incubated on TY agar lacking IPTG and ~20% of transconjugants carried the construct integrated into the genome (Figure S15b). Induction of Cre expression and simultaneous activation of *oriV*$_{pMlb\_lacO}$ by IPTG resulted in excision and maintenance of the floxed DNA fragment, respectively, giving rise to strain *S. meliloti* SmCre_exo-OUT carrying pIso-exo as autonomous mini-replicon. Successful Cre-mediated recombination, resulting in an active *loxP* site on pIso-exo and *loxLR* remaining on pSymB, was verified via sequencing (Figure S15c). Since the recombination event caused an antibiotic resistance switch, the screening effort for identifying positive clones was minimized. ABC cloning provided high efficiency rates of ~75 % after Cre/*lox*-mediated isolation, which can be attributed to highly efficient Cre recombination in SmCre derivatives[45] in general, and suitable function of the engineered *repABC*-type *oriV* in particular. Since multiple *lox* sites recombine in an unpredictable manner, a single copy number of activated *repABC*$_{pMlb}$-*lacO* plasmids is an important prerequisite for manageable Cre reactions. Promoter modification appeared to affect propagation stability of pMlb_*lacO* and pIso-exo, while retaining copy number control, as suggested by qPCR-based copy number determination (Figure S15d).

Transfer of pIso-exo to *E. coli* was achieved via triparental mating. $oriV_{pMB1}$ on the pIso backbone enabled replication in *E. coli* and due to the high copy number efficient purification of the large DNA molecule from this host (Figure 7b ii). Moreover, conjugation-mediated transfer to *A. tumefaciens* C58 exoB resulted in complementation of the exopolysaccharide-deficient phenotype (Figure 7b iii), demonstrating propagation of the engineered *repABC* cassette also in this host (Figure S15e). Thus, *repABC*-based cloning approaches appear applicable in a much wider range of α-rhizobia.

Conclusions. The currently available ectopic expression systems applicable in *S. meliloti* and related bacteria are based on a small number of low- and multi-copy vectors. Due to incompatibilities, options for combining different vectors in one cell are very limited. Moreover, different copy number ranges and cell-to-cell variation complicate balanced gene expression even from a set of compatible plasmids. While chromosomal integration of expression constructs solves copy number problems, stable integration of multiple constructs is still cumbersome and integration sites may influence gene expression. In this study, we established the pABC vector family which circumvents these issues and highly facilitates the setup of combinatorial expression systems.

pABC vectors provide a well-defined integration site in single copy and we demonstrated stable co-existence of up to three compatible pABCs. Occurrence of α-rhizobial species harboring up to six RepABC-family plasmids[24] suggests that even more than three pABC plasmids may be combined in a suitable host. The number of available resistance cassettes and restraints in the applicability of multiple antibiotics selection pressure potentially hampers combinatorial use of pABC vectors. However, $oriV_{repABC}$-based multi-replicon systems benefit from robust inheritance under non-selective conditions. A future route to further stabilization of pABC plasmids, independent of antibiotic resistance cassettes, could be the integration of toxin-antitoxin (TA)-based maintenance systems, which frequently occur in *S. meliloti*[67]. Integration of additional modules is facilitated by the modular pABC design, which also facilitates construction of pABCs with multifarious part combinations. pABC plasmids were designed as shuttle vectors capable of replication in *E. coli* and *S. meliloti*. This is an advantageous feature when it comes to *E. coli*-based cloning procedures and plasmid DNA purification. Nevertheless, some

applications in *S. meliloti* may favor pABC plasmids lacking oriVEc and oriT. We demonstrated that these modules can be efficiently removed on demand by induction of Cre/*lox*-mediated site specific recombination in *S. meliloti*.

Furthermore, we took a stab at the broad potential of pABC applications by devising the *in vivo* ABC cloning procedure, which provides high flexibility due to isolation and immediate maintenance of the fragment of interest on a *repABC*-based vector. This opens up multifarious follow-up processes including intracellular relocation of the isolated fragment by Cre/*lox* mediated recombination or conjugal transfer to a compatible host organism. In a future perspective, the available set of non-promiscuous *lox* spacer mutants[45] paves the way to serial engineering steps, such as *in vivo* assembly of indigenous or heterologous DNA fragments on pABC plasmids. Thus, the family of pABC vectors has the potential to significantly advance metabolic and genome engineering in *S. meliloti* and related species and to provide a platform for *in vivo* DNA assembly.

■ METHODS

Bacterial strains and growth conditions. Bacterial strains used in this study are derivatives of *Escherichia coli* K12, *Sinorhizobium meliloti* Rm1021, *Agrobacterium tumefaciens* C58, *Rhizobium etli* CFN42, *Sinorhizobium fredii* NGR234, *Mesorhizobium loti* MAFF303099, and *Sinorhizobium medicae* WSM419. *E. coli* strains were cultivated at 37°C, whereas *A. tumefaciens*, *M. loti*, *R. etli*, *S. fredii and S. medicae* were grown at 30°C in Luria-Bertani (LB) medium[68]. *S. meliloti* strains were cultivated at 30°C in tryptone yeast extract (TY) medium[69]. If necessary, the following antibiotics were used: Gentamicin (8 µg/ml for *E. coli*, 30 µg/ml for *S. meliloti*, and 50µg/ml for *A. tumefaciens*), kanamycin (50 µg/ml for *E. coli* and 200 µg/ml for *S. meliloti*), streptomycin (600 µg/ml for *S. meliloti*), spectinomycin (100 µg/ml for *E. coli*, 200 µg/ml for *S. meliloti*, and 500 µg/ml for *A. tumefaciens*), hygromycin (100 µg/ml for both *E. coli* and *S. meliloti*), trimethoprim (1000 µg/ml for *E. coli*, 700 µg/ml for *S. meliloti*), chloramphenicol (20 µg/ml for *E. coli*) or tetracycline (10 µg/ml for *E. coli* and 3 µg/ml for *S. meliloti*). Cre/*lox* applications were performed as described before[46], excluding tetracycline supplementation during

induction of *cre* expression. Calcofluor fluorescence assays of *A. tumefaciens* C58 exoB carrying pIso-exo were performed on LB agar supplemented with gentamycin and spectinomycin, 200 µg/ml calcofluor and 500 µM IPTG. Cultures of an optical density at 600 nm (OD_{600}) of 0.2 were spotted on agar medium and grown overnight at 30°C.

DNA manipulation and Plasmid extraction. Standard techniques were employed for common cloning procedures[68]. Primers and PCR products were 5'-phosphorylated applying T4 Polynucleotide Kinase (Thermo Scientific, Germany). Dephosphorylation of 5'-ends of DNA was achieved by use of FastAP Thermosensitive Alkaline Phosphatase (Fermentas). Blunting of DNA ends was performed by use of DNA Polymerase I (Large Klenow fragment) (New England Biolabs (U.K.)). Plasmid DNA from *E. coli* and *S. meliloti* strains was purified using the "E.Z.N.A. Plasmid Mini Kit" (Omega Bio-Tek). For gel extractions of DNA the "illustra GFX PCR DNA and Gel Band Purification Kit" (GE Healthcare Life Sciences) was employed.

Plasmid and strain construction. Strains and constructs generated in this study are listed in Tables 1 and S9. Assembled pABCs are listed in Table 2. A detailed description of plasmids is given in Tables S7 and S10. Oligonucleotides and synthetic DNA fragments were provided by Sigma-Aldrich (USA) and Integrated DNA Technologies (USA), respectively (Table S2). DNA fragments were PCR amplified using Q5® High-Fidelity DNA Polymerase (New England Biolabs) or Taq DNA Polymerase (New England Biolabs). DNA sequence verification was performed by Eurofins Genomics (Germany). For transformation of *S. meliloti*, the respective strains were either conjugated or electroporated as described before[45]. Preparation of electrocompetent cells was done as described by Ferri et al. (2010).

Construction of pABC mini-replicons: Amplification of module parts from library plasmids was performed with Q5 polymerase applying module specific primer sets 494/495 (pLoriVSm derivatives), 496/497 (pLsynTer derivatives), 500/501 (pLoriT derivatives) and 502/503 (pLAR derivatives). oriVEc module parts were amplified from pK18mob2 (pMB1), pACYC184 (p15A) and pX-G10 (pSC101*) with primer sets 546/547, 544/545 and 670/671, respectively. Subsequently, matching DNA fragments (blunt-end and 5' phosphorylated) were fused via ligase cycling reaction (LCR) according to de Kok et al.

(2014). For LCR 0.3 U Taq ligase (New England Biolabs), 3 nM DNA parts, 30 nM bridging oligos and 8% DMSO were applied. The following conditions were used: 2 min at 94°C and then 50 cycles of 10 s at 94°C, 30 s at 60°C, and 60 s at 65°C, followed by incubation at 4°C. Reaction mixtures were purified using the E.Z.N.A. Cycle-Pure Kit (Omega Bio-Tek).

For live-cell localization studies of pABC vectors, a dual label system based on the fluorescent reporter gene fusions *tetR-eYFP* (derived from the fluorescent repressor operator system, FROS[58]) and *mChr-parB*$_{Yp}$ (*parB* derived from *Yersinia pestis*) was developed. Homologous regions comprising the upstream region of *tauA* (Smb1050453-1050952) and P*tauAB-tauA* (Smb1050781-1051458) were integrated into pK18mobsacB, thereby flanking *tetR-YFP-mChr-parB*$_{Yp}$. *S. meliloti* Rm1021 was transformed with the resulting plasmid pJD120. Loss of the plasmid backbone was forced by sucrose selection resulting in strain JDSm106 carrying the *tetR-YFP-mChr-parB*$_{Yp}$ reporter construct stably integrated into the chromosome downstream of the *tauA* promoter. Localization studies of pABCs carrying either a pLau44-derived *tetO* array[58] (position 5117-9290) or a pMT1-derived, 86 bp *parS*$_{Yp}$ sequence[70] were performed in this dual label *S. meliloti* reporter strain JDSm106.

Control strains for cell growth and plasmid inheritance experiments were generated as follows. A homologous region comprising *exsA* (Smb1166189-1166773) was cloned into pK18mob2, resulting in pMW1. Integration of a P*min2-eGFP* cassette upstream of the *exsA* region in pMW1 gave rise to pMW28. pMW1 was integrated into pSymB in *S. meliloti* Rm1021 by homologous recombination, resulting in the kanamycin resistant strain MWSm8. Integration of pMW28 into the genome of *S. meliloti* ΔhsdR by homologous recombination gave rise to strain MWSm37 exhibiting constitutive *eGFP* expression.

An *S. meliloti* strain suitable for ABC cloning of the *exo* gene cluster was constructed in a two-step cloning strategy applying cloning-free genome editing (CFGE)[45]. P*min2-loxR-aacC1* (PCR amplified from pJD211 with primers Uni/Rev) was blunt-end ligated with a ~550 bp fragment covering the *exoP-thiD* intergenic region (Smb1190665-1191220; PCR amplified with primers 86/87). Purified ligation products were then introduced to SmCreΔhsdR by electroporation. Finally, a transformant exhibiting

P*min2-loxR-aacC1* correctly integrated into the genome downstream of the *exo* gene cluster was transformed with pIso, giving rise to strain SmCre_exo-IN.

Transfer of plasmid DNA from *S. meliloti* to *E. coli* and *A. tumefaciens*. In order to transfer pISO-exo from *S. meliloti* into different species, triparental matings between *S. meliloti* SmCre_exo-OUT and *E. coli* XL1blue-pRK2013, with *A. tumefaciens* C58 exoB or *E. coli* AB3219 were performed.

The cell density of fresh overnight cultures was adjusted to an OD_{600} of 2.0 and equal volumes of donor, recipient and helper strain were mixed. 50 µl of cell suspension were dropped on LB agar supplemented with 500 µM IPTG and incubated overnight. Cells were recovered in 1 ml sterile 0.9% NaCl solution. For the determination of conjugation efficiency, OD_{600} of the cell suspensions was adjusted to 2.0 and appropriate dilutions were plated on selective LB agar containing 500 µM IPTG. The efficiency of conjugation was determined as the number of colony forming units (*cfu*) per ml suspension.

Plasmid Stability Assays. In order to assess inheritance stability of pLoriVSm derivatives, *S. meliloti* cultures were grown over 72 h at 30°C in TY medium supplemented with streptomycin. Cell suspensions were diluted every 12 h to an OD_{600} of ~0.05. Finally, dilution series were plated on non-selective TY agar and emerging colonies were streaked on both selective (Km) and non-selective TY agar. Antibiotic resistance was correlated with the presence of the assessed plasmid.

For determination of propagation stability of pABC-*eGFP* derivatives, *S. meliloti* cultures were incubated over 96 h at 30°C in TY medium. Probes were taken 0, 24, 48 and 96 h post-inoculation and immediately diluted to an OD_{600} of 0.1 in 0.9% NaCl. eGFP fluorescence was detected using a BD Fortessa flow cytometer. For FACS analysis 10.000 gated events were acquired giving rise to the ratio between eGFP-negative and positive cells. Sufficient sensitivity was verified (Figure S9).

qPCR based copy number determination. *S. meliloti* cultures were grown in selective TY medium until exponential growth phase was reached and adjusted to ~$1.39x10^5$ cells/µl (≙ 1 ng DNA/µl). Cell suspensions that had been incubated at 95°C for 15 minutes served as template. The DNA concentration of cell lysates containing pLoriVSm derivatives was verified via Qubit™ fluorometric quantitation (Thermo Fisher Scientific). qRT-PCR was carried out in a qTOWER Thermal Cycler

(Analytik Jena, Germany) using the Taykon™No Rox SYBR®MasterMix dTTP Blue Kit. Reactions were performed according to the manufacturer's instructions in a 5 µl reaction volume. Replicon copy number was calculated according to Lee et al. (2006)[71].

Microscopy. Phase contrast and fluorescence pictures of *S. meliloti* were taken using an Eclipse Ti-E inverse research microscope (Nikon) and the live imaging modus of the NIS elements software (Nikon). For fluorescence microscopy with a 100x CFI Apo TIRF Oil objective (numerical aperture of 1.49) a green DPSS solid state laser (561 nm, 50 mW; Sapphire) and a multiline Argonlaser (65 mW; Melles Griot) were applied with AHF HC filter sets F36-504 TxRed (excitation bp 562/40 nm, beam splitter 593 nm, emission bp 624/40 nm) and F36-513 YFP (excitation bp 500/24 nm, beam splitter 520 nm, emission bp 542/27 nm). Background fluorescence was reduced by applying the highly inclined laminated optical sheet (HILO) technique[72]. Cell cultures were grown in TY medium supplemented with suitable antibiotics to an OD_{600} of 0.1 and expression of the fluorescent labeling system was induced with 15 mM taurine for 4h. Living cells were placed onto 1% TY agarose pads.

Bioinformatics tools. In order to generate neutral orthogonal DNA sequences for module linker fragments, the web-based tool R2oDNA Designer[73] was applied. To estimate the recombination potential of distinct pABCs we investigated the local sequence homology of all module parts (including linker sequences). We applied a window based local alignment approach to focus on the local sequence homology. For a pair of parts we therefore extracted all 200 bp sequence fragments of one part and performed a local alignment on the full sequence of the second part. The best alignment similarity of all 200 bp fragments with the second vector was taken as the similarity value of both vectors.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +49 (621) 28-24451. E-mail: anke.becker@synmikro.uni-marburg.de

Author Contributions

A.B., J.D., and M.W. designed the research. J.D., M.W., and C.H. performed the experiments. P.S. computationally analyzed parts for sequence similarities. J.D. and M.W analyzed all data. J.D. and A.B. wrote the paper.

Notes

The authors declare no competing financial interest.

Acknowledgment

Financial support from the German Research Foundation (Collaborative Research Centre TRR 174 “Spatiotemporal dynamics of bacterial cells”), and the LOEWE program of the State of Hesse (Germany) is acknowledged.

REFERENCES

1. Rajewska, M., Wegrzyn, K., and Konieczny, I. (2012) AT-rich region and repeated sequences - the essential elements of replication origins of bacterial replicons. *FEMS Microbiol. Rev. 36*, 408–434.
2. Gerdes, K., Howard, M., and Szardenings, F. (2010) Pushing and pulling in prokaryotic DNA segregation. *Cell 141*, 927–942.
3. Martinez-Garcia, E., Benedetti, I., Hueso, A., and Lorenzo, V. de (2015) Mining Environmental Plasmids for Synthetic Biology Parts and Devices. *Microbiol. Spectrum 3*, PLAS-0033-2014.
4. Messerschmidt, S. J., Kemter, F. S., Schindler, D., and Waldminghaus, T. (2015) Synthetic secondary chromosomes in *Escherichia coli* based on the replication origin of chromosome II in *Vibrio cholerae*. *Biotechnol. J. 10*, 302–314.
5. Milbredt, S., Farmani, N., Sobetzko, P., and Waldminghaus, T. (2016) DNA Replication in Engineered *Escherichia coli* Genomes with Extra Replication Origins. *ACS Synth. Biol.* Epub Jun 17, 2016. DOI: 10.1021/acssynbio.6b00064.
6. Liang, X., Baek, C.-H., and Katzen, F. (2013) *Escherichia coli* with two linear chromosomes. *ACS Synth. Biol. 2*, 734–740.
7. Itaya, M. (1999) Effective Cloning of Unmarked DNA Fragments in the *Bacillus subtilis* 168 Genome. *Biosci., Biotechnol., Biochem. 63*, 602–604.
8. Itaya, M., Tsuge, K., Koizumi, M., and Fujita, K. (2005) Combining two genomes in one cell: stable cloning of the *Synechocystis* PCC6803 genome in the *Bacillus subtilis* 168 genome. *Proc. Natl. Acad. Sci. U. S. A. 102*, 15971–15976.
9. Kaneko, S., Akioka, M., Tsuge, K., and Itaya, M. (2005) DNA shuttling between plasmid vectors and a genome vector: systematic conversion and preservation of DNA libraries using the *Bacillus subtilis* genome (BGM) vector. *J. Mol. Biol. 349*, 1036–1044.
10. MacLean, A. M., Finan, T. M., and Sadowsky, M. J. (2007) Genomes of the symbiotic nitrogen-fixing bacteria of legumes. *Plant Physiol. 144*, 615–622.
11. Venturi, V., and Keel, C. (2016) Signaling in the Rhizosphere. *Trends Plant Sci. 21*, 187–198.
12. Zhu, J., Oger, P. M., Schrammeijer, B., Hooykaas, P. J. J., Farrand, S. K., and Winans, S. C. (2000) The Bases of Crown Gall Tumorigenesis. *J. Bacteriol. 182*, 3885–3895.
13. Smith, W. F., and Townsend, C. O. (1907) A Plant-Tumor of Bacterial Origin. *Science 25*, 671-673.
14. Mus, F., Crook, M.B., Garcia, K., Garcia Costas, A., Geddes, B. A., Kouri, E. D., Paramasivan, P., Ryu, M.-H., Oldroyd, G. E. D., Poole, P. S., Udvardi, M. K., Voigt, C.A., Ané, J.-M., and Peters, J. W. (2016) Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes. *Appl. Environ. Microbiol. 82*, 3698–3710.
15. Castellane, T. C. L., Lemos, M. V. F., and Lemos, E. G. de M. (2014) Evaluation of the biotechnological potential of *Rhizobium tropici* strains for exopolysaccharide production. *Carbohydr. Polym. 111*, 191–197.

16. Wu, D., Li, A., Ma, F., Yang, J., and Xie, Y. (2016) Genetic control and regulatory mechanisms of succinoglycan and curdlan biosynthesis in genus *Agrobacterium*. *Appl. Microbiol. Biotechnol. 100*, 6183–6192.

17. Tazoe, M., Ichikawa, K., and Hoshino, T. (1999) Production of Vitamin B6 in *Rhizobium*. *Biosci., Biotechnol., Biochem. 63*, 1378–1382.

18. Dong, H., Li, S., Fang, H., Xia, M., Zheng, P., Zhang, D., and Sun, J. (2016) A newly isolated and identified vitamin B12 producing strain: *Sinorhizobium meliloti* 320. *Bioprocess Biosyst. Eng. 39*, 1527–1537.

19. Gelvin, S. B. (2003) *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol. Mol. Biol. Rev. 67*, 16-37.

20. Harrison, P. W., Lower, R. P., Kim, N. K., and Young, J. P. (2010) Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol. 18*, 141–148.

21. Mazur, A., and Koper, P. (2012) Rhizobial plasmids — replication, structure and biological role. *Cent. Eur. J. Biol. 7*, 571-586.

22. Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., and Tabata, S. (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res. 7*, 331–338.

23. Galibert, F., Finan, T.A., Long, S.R., Pühler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M. J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., Dreano, S., Federspiel, N. A., Fisher, R. F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R. W., Jones, T., Kahn, D., Kahn, M.L., Kalman, S., Keating, D. H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M. C., Pohl, T. M., Portetelle, D., Purnelle, B., Ramsperger, B. U., Surzycki, R., Thébault, P., Vandenbol, M., Vorhölter, F.-J., Weidner, S., Wells, D. H., Wong, K., Yeh, K.-C., and Batut, J. (2001) The Composite Genome of the Legume Symbiont *Sinorhizobium meliloti*. *Science 293*, 668–672.

24. González, V., Santamaría, R. I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S. C., Ramírez, M. A., Jiménez-Jacinto, V., Collado-Vides, J., and Dávila, G. (2006) The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc. Natl. Acad. Sci. U. S. A. 103*, 3834–3839.

25. Wood, D. W., Setubal, J. C., Kaul, R., Monks, D. E., Kitajima, J. P., Okura, V. K., Zhou, Y., Chen, L., Wood, G. E., Almeida, N. F. Jr, Woo, L., Chen, Y., Paulsen, I. T., Eisen, J. A., Karp, P. D., Bovee, D., SR, Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Kutyavin, T., Levy, R., Li, M. J., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Romero, P., Gordon, D., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z. Y., Dolan, M., Chumley, F., Tingey, S. V., Tomb, J. F., Gordon, M. P., Olson, M. V., and Nester, E. W. (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science 294*, 2317–2323.

26. Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman, B. S., Cao, Y., Askenazi, M., Halling, C., Mullin, L., Houmiel, K., Gordon, J., Vaudin, M., Iartchouk, O., Epp, A., Liu, F., Wollam, C., Allinger, M., Doughty, D., Scott, C., Lappas, C., Markelz, B., Flanagan, C., Crowell, C., Gurson, J., Lomo, C., Sear, C., Strub, G., Cielo, C., and Slater, S. (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science 294*, 2323–2328.

27. Frage, B., Döhlemann, J., Robledo, M., Lucena, D., Sobetzko, P., Graumann, P. L., and Becker, A. (2016) Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. *Mol. Microbiol. 100*, 808–823.

28. Pinto, U. M., Pappas, K. M., and Winans, S. C. (2012) The ABCs of plasmid replication and segregation. *Nat. Rev. Microbiol. 10*, 755–765.

29. Cevallos, M. A., Cervantes-Rivera, R., and Gutiérrez-Ríos, R. M. (2008) The *repABC* plasmid family. *Plasmid 60*, 19–37.

30. Brom, S., García de los Santos, A., Stepkowsky, T., Flores, M., Dávila, G., Romero, D., and Palacios, R. (1992) Different Plasmids of *Rhizobium leguminosarum* bv. *phaseoli* Are Required for Optimal Symbiotic Performance. *J. Bacteriol. 174*, 5183–5189.

31. Brom, S., García-de los Santos, A., Cervantes, L., Palacios, R., and Romero, D. (2000) In *Rhizobium etli* symbiotic plasmid transfer, nodulation competitivity and cellular growth require interaction among different replicons. *Plasmid 44*, 34–43.

32. Landeta, C., Dávalos, A., Cevallos, M. Á., Geiger, O., Brom, S., and Romero, D. (2011) Plasmids with a chromosome-like role in rhizobia. *J. Bacteriol. 193*, 1317–1326.

33. Chai, Y., and Winans, S. C. (2005) RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between *repA* and *repB* for plasmid partitioning and autorepression. *Mol. Microbiol. 58*, 1114–1129.

34. Żebracki, K., Koper, P., Marczak, M., Skorupska, A., and Mazur, A. (2015) Plasmid-Encoded RepA Proteins Specifically Autorepress Individual *repABC* Operons in the Multipartite *Rhizobium leguminosarum* bv. *trifolii* Genome. *PLoS One 10*, e0131907.

35. Cervantes-Rivera, R., Romero-López, C., Berzal-Herranz, A., and Cevallos, M. A. (2010) Analysis of the mechanism of action of the antisense RNA that controls the replication of the *repABC* plasmid p42d. *J. Bacteriol. 192*, 3268–3278.

36. Rivera-Urbalejo, A., Pérez-Oseguera, Á., Carreón-Rodríguez, O. E., and Cevallos, M.A. (2015) Mutations in an antisense RNA, involved in the replication control of a *repABC* plasmid, that disrupt plasmid incompatibility and mediate plasmid speciation. *Plasmid 78*, 48–58.

37. Pappas, K. M., and Winans, S. C. (2003) A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. *Mol. Microbiol. 48*, 1059–1073.

38. McAnulla, C., Edwards, A., Sanchez-Contreras, M., Sawers, R. G., and Downie, J. A. (2007) Quorum-sensing-regulated transcriptional initiation of plasmid transfer and replication genes in *Rhizobium leguminosarum biovar viciae*. *Microbiology 153*, 2074–2082.

39. Pérez-Oseguera, A., and Cevallos, M. A. (2013) RepA and RepB exert plasmid incompatibility repressing the transcription of the *repABC* operon. *Plasmid 70*, 362–376.

40. Yip, C. B., Ding, H., and Hynes, M. F. (2015) Counter-transcribed RNAs of *Rhizobium leguminosarum repABC* plasmids exert incompatibility effects only when highly expressed. *Plasmid 78*, 37–47.

41. Koper, P., Żebracki, K., Marczak, M., Skorupska, A., and Mazur, A. (2016) RepB proteins of the multipartite *Rhizobium leguminosarum* bv. *trifolii* genome discriminate between centromere-like *parS* sequences for plasmid segregational stability. *Mol. Microbiol.* Epub Aug 2, 2016. DOI: 10.1111/mmi.13472.

42. Ding, H., and Hynes, M. F. (2009) Plasmid transfer systems in the rhizobia. *Can. J. Microbiol. 55*, 917–927.

43. Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and Pühler, A. (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene 145*, 69–73.

44. Becker, A., Barnett, M. J., Capela, D., Dondrup, M., Kamp, P.-B., Krol, E., Linke, B., Rüberg, S., Runte, K., Schroeder, B. K., Weidner, S., Yurgel, S. N., Batut, J., Long, S. R., Pühler, A., and Goesmann, A. (2009) A portal for rhizobial genomes: RhizoGATE integrates a *Sinorhizobium meliloti* genome annotation update with postgenome data. *J. Biotechnol.140*, 45–50.

45. Döhlemann, J., Brennecke, M., and Becker, A. (2016) Cloning-free genome engineering in *Sinorhizobium meliloti* advances applications of Cre/*loxP* site-specific recombination. *J. Biotechnol.10*, 160-170.

46. Harrison, C. L., Crook, M. B., Peco, G., Long, S. R., and Griffitts, J. S. (2011) Employing site-specific recombination for conditional genetic analysis in *Sinorhizobium meliloti*. *Appl. Environ. Microbiol. 77*, 3916–3922.

47. de Kok, S., Stanton, L. H., Slaby, T., Durot, M., Holmes, V. F., Patel, K. G., Platt, D., Shapland, E. B., Serber, Z., Dean, J., Newman, J. D., and Chandran, S. S. (2014) Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synth. Biol. 3*, 97–106.

48. Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de las Heras, A., Páez-Espino, A. D., Durante-Rodríguez, G., Kim, J., Nikel, P. I., Platero, R., and de Lorenzo, V. (2012) The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids. Res. 41*, D666-675.

49. Datta, N., Hedges, R. W., Shaw, E. J., Sykes, R. B., and Richmond, M. H. (1971) Properties of an R Factor from *Pseudomonas aeruginosa*. *J. Bacteriol. 108*, 1244–1249.

50. Li, P.L., and Farrand, S.K. (2000) The Replicator of the Nopaline-Type Ti Plasmid pTiC58 Is a Member of the *repABC* Family and Is Influenced by the TraR-Dependent Quorum-Sensing Regulatory System. *J. Bacteriol. 182*, 179–188.

51. Charoenpanich, P., Soto, M.J., Becker, A., and McIntosh, M. (2015) Quorum sensing restrains growth and is rapidly inactivated during domestication of *Sinorhizobium meliloti*. *Environ. Microbiol. Rep. 7*, 373–382.

52. Pellock, B. J., Teplitski, M., Boinay, R. P., Bauer, W. D., and Walker, G. C. (2002) A LuxR Homolog Controls Production of Symbiotically Active Extracellular Polysaccharide II by *Sinorhizobium meliloti*. *J. Bacteriol. 184*, 5067–5076.

53. Gao, M., Chen, H., Eberhard, A., Gronquist, M. R., Robinson, J. B., Rolfe, B. G., and Bauer, W. D. (2005) *sinI*- and *expR*-dependent quorum sensing in *Sinorhizobium meliloti*. *J. Bacteriol. 187*, 7931–7944.

54. Hoang, H. H., Becker, A., and González, J. E. (2004) The LuxR homolog ExpR, in combination with the Sin quorum sensing system, plays a central role in *Sinorhizobium meliloti* gene expression. *J. Bacteriol. 186*, 5460–5472.

55. Glenn, S. A., Gurich, N., Feeney, M. A., and González, J. E. (2007) The ExpR/Sin quorum-sensing system controls succinoglycan production in *Sinorhizobium meliloti*. *J. Bacteriol. 189*, 7077–7088.

56. Marketon, M. M., Glenn, S. A., Eberhard, A., and Gonzalez, J. E. (2003) Quorum Sensing Controls Exopolysaccharide Production in *Sinorhizobium meliloti*. *J. Bacteriol. 185*, 325–331.

57. Schwartz, M. A., and Shapiro, L. (2011) An SMC ATPase mutant disrupts chromosome segregation in *Caulobacter*. *Mol. Microbiol. 82*, 1359–1374.

58. Lau, I. F., Filipe, S. R., Søballe, B., Økstad, O.-A., Barre, F.-X., and Sherratt, D. J. (2003) Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol. Microbiol. 49*, 731–743.

59. Cocchia, M., Kouprina, N., Kim, S.-J., Larionov, V., Schlessinger, D., and Nagaraja, R. (2000) Recovery and potential utility of YACs as circular YACs/BACs. *Nucleic Acids Res. 28*, e81.

60. Leem, S.-H., Noskov, V.N., Park, J.-E., Kim, S. I., Larionov, V., and Kouprina, N. (2003) Optimum conditions for selective isolation of genes from complex genomes by transformation-associated recombination cloning. *Nucleic Acids Res. 31*, e29.

61. Feng, Z., Kim, J. H., and Brady, S. F. (2010) Fluostatins produced by the heterologous expression of a TAR reassembled environmental DNA derived type II PKS gene cluster. *J. Am. Chem. Soc. 132*, 11902–11903.

62. Shao, Z., Luo, Y., Zhao, H. (2011) Rapid characterization and engineering of natural product biosynthetic pathways via DNA assembler. *Mol. Biosyst. 7*, 1056–1059.

63. Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P. M., Plaza, A., Xia, L., Müller, R., Stewart, A. F., and Zhang, Y. (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat. Biotechnol. 30*, 440–446.

64. Du, D., Wang, L., Tian, Y., Liu, H., Tan, H., and Niu, G. (2015) Genome engineering and direct cloning of antibiotic gene clusters via phage ϕBT1 integrase-mediated site-specific recombination in Streptomyces. *Sci. Rep. 5*, 8740.

65. Hu, S., Liu, Z., Zhang, X., Zhang, G., Xie, Y., Ding, X., Mo, X., Stewart, A. F., Fu, J., Zhang, Y., and Xia, L. (2016) "Cre/*loxP* plus BAC": a strategy for direct cloning of large DNA fragment and its applications in *Photorhabdus luminescens* and *Agrobacterium tumefaciens*. *Sci. Rep. 6*, 29087.

66. Chain, P. S., Hernandez-Lucas, I., Golding, B., and Finan, T. M. (2000) *oriT*-Directed Cloning of Defined Large Regions from Bacterial Genomes: Identification of the *Sinorhizobium meliloti* pExo Megaplasmid Replicator Region. *J. Bacteriol. 182*, 5486–5494.

67. Milunovic, B., diCenzo, G. C., Morton, R. A., and Finan, T. M. (2014) Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of *Sinorhizobium meliloti*. *J. Bacteriol. 196*, 811–824.

68. Green, M. R., and Sambrook, J. (2012) *Molecular cloning: A laboratory manual*, 4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

69. Beringer, J. E. (1974) R Factor Transfer in *Rhizobium leguminosarum*. *J. Gen. Microbiol. 84*, 188–198.

70. Lindler, L. E., Plano, G. V., Burland, V., Mayhew, G. F., and Blattner, F. R.(1998) Complete DNA Sequence and Detailed Analysis of the *Yersinia pestis* KIM5 Plasmid Encoding Murine Toxin and Capsular Antigen. *Infect. Immun. 66*, 5731–5742.

71. Lee, C., Kim, J., Shin, S. G., and Hwang, S. (2006) Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. *J. Biotechnol. 123*, 273–280.

72. Tokunaga, M., Imamoto, N., and Sakata-Sogawa, K. (2008) Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat. Methods 5*, 159–161.

73. Casini, A., Christodoulou, G., Freemont, P. S., Baldwin, G. S., Ellis, T., and MacDonald, J. T. (2014) R2oDNA designer: computational design of biologically neutral synthetic DNA sequences. *ACS Synth. Biol. 3*, 525–528.

74. Cervantes-Rivera, R., Pedraza-López, F., Pérez-Segura, G., and Cevallos, M. A. (2011) The replication origin of a *repABC* plasmid. *BMC Microbiol. 11*, 158.

75. Bignell, C., and Thomas, C. M. (2001) The bacterial ParA-ParB partitioning proteins. *J. Biotechnol. 91*, 1–34.

76. Quinto, C., de la Vega, H., Flores, M., Fernández, L., Ballado, T., Soberón, G., and Palacios, R. (1982) Reiteration of nitrogen fixation gene sequences in *Rhizobium phaseoli*. *Nature 299*, 724–726.

77. Reeve, W., Chain, P., O'Hara, G., Ardley, J., Nandesena, K., Bräu, L., Tiwari, R., Malfatti, S., Kiss, H., Lapidus, A., Copeland, A., Nolan, M., Land, M., Hauser, L., Chang, Y.-J., Ivanova, N., Mavromatis, K., Markowitz, V., Kyrpides, N., Gollagher, M., Yates, R., Dilworth, M., and Howieson, J. (2010) Complete genome sequence of the *Medicago* microsymbiont *Ensifer* (*Sinorhizobium*) *medicae* strain WSM419. *Stand. Genomic Sci.2*, 77–86.

78. Perret, X., Fellay, R., Bjourson, A. J., Cooper, J. E., Brenner, S., and Broughton, W. J. (1994) Subtraction hybridisation and shot-gun sequencing: a new approach to identify symbiotic loci. *Nucleic Acids Res. 22*, 1335–1341.

79. Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol. 166*, 557–580.

80. Simon, R., Priefer, U., and Pühler, A. (1983) A Broad Host Range Mobilization System for *In Vivo* Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Nat. Biotech. 1*, 784–791.

81. Zähringer, F., Lacanna, E., Jenal, U., Schirmer, T., and Boehm, A. (2013) Structure and signaling mechanism of a zinc-sensory diguanylate cyclase. *Structure 21*, 1149–1157.

82. Khan, S. R., Gaines, J., Roop, R. M. 2nd, and Farrand, S. K. (2008) Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl. Environ. Microbiol. 74*, 5053–5062.

83. Urban, J. H., and Vogel, J. (2007) Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. *Nucleic Acids Res. 35*, 1018–1037.

84. Chang, A. C. Y., and Cohen, S. N. (1978) Construction and Characterization of Amplifiable Multicopy DNA Cloning Vehicles Derived from the P15A Cryptic Miniplasmid. *J. Bacteriol. 134*, 1141–1156.

85. Hübner, P., Willison, J. C., Vignais, P. M., and Bickle, T. A. (1991) Expression of regulatory *nif* genes in *Rhodobacter capsulatus*. *J. Bacteriol. 173*, 2993–2999.

86. Figurski, D.H., and Helinski, D.R. (1979). Replication of an origin-containing derivative of RK2 dependent on a plasmid function provided *in trans*. *Proc. Natl. Acad. Sci. U. S. A*. 76, 1648-1652.

Supporting Information

Supplementary tables

Supplementary figures

Abbreviations

Clm: chloramphenicol

Gm: gentamicin

h: hour

Hyg: hygromicin

Km: kanamycin

oriV: replication origin

SD: Shine-Dalgarno sequence

Sp: spectinomycin

Str: streptomycin

TA: toxin-antitoxin

Tc: tetracycline

Tmp: trimethoprim

Table 1. Bacterial strains and plasmids used in this study.

Strains/ Plasmids	Description	References
S. meliloti		
Rm1021	wild type strain	Galibert et al.,2001
ΔhsdR	Rm1021 *cre* expression strain with *hsdR* deletion	Döhlemann et al., 2016
SmCreΔhsdR	Rm1021 *cre* expression strain with *hsdR* deletion, *tauX*:: *cre-tetRA* (Tc^r)	Döhlemann et al., 2016
JDSm106	Rm1021 bearing *tetR-YFP-mChr-parB*$_{Yp}$ (pJD120) under control of P*tauAB*	This study
MWSm8	Rm1021 bearing pMW1 genomically integrated (Km^r)	This study
MWSm37	ΔhsdR bearing pMW28 for constitutive *eGFP* expression (Km^r)	This study
SmCre_exo-IN	SmCreΔhsdR bearing pIso and P*min2-loxR-aacC1* integrated up- and downstream of the *exo* gene cluster, respectively (Km^r, Gm^r)	This study
SmCre_exo-OUT	SmCreΔhsdR carrying pIso-exo isolated from pSymB (IPTG dependent) (Km^r, $Spec^r$)	This study
A. tumefaciens		
C58	wild type strain	Goodner et al., 2001; Wood et al., 2001
C58 exoB	exhibiting a CFGE-mediated inactivation of *exoB* (Gm^r)	Döhlemann et al. 2016
exoB/pIso-exo	C58 exoB carrying pIso-exo (IPTG dependent) (Gm^r, Km^r, $Spec^r$)	This study
R. etli		
CFN42	wild type strain	Quinto et al.,1982
S. medicae		
WSM419	wild type strain	Reeve et al., 2010
S. fredii		
NGR234	wild type strain	Perret et al., 1994
M. loti		
MAFF303099	wild type strain	Kaneko et al., 2000
E. coli		
DH5α	used for cloning procedures and plasmid extraction	Hanahan, 1983
S17-1	used for conjugation with *S. meliloti*	Simon et al., 1983
XL1-Blue	used as recipient of pRK2013 (Tc^r)	Stratagene
XL1Blue-pRK2013	used for triparental matings (Tc^r, Km^r)	This study
AB3219	used as Clm resistant recipient of pISO-exo (Clm^r)	Zähringer et al., 2013
Plasmids		
pK/G18mob2	suicide vector carrying replication origin pMB1	Schäfer et al., 1994
pK18mobsacB	suicide vector carrying *sacB* for sucrose selection	Schäfer et al., 1994
pSRK-Gm/Km	broad-host-range expression vector	Khan et al., 2008
pXG-10	carrying replication origin pSC101* (derivative of pSC101 exhibiting reduced copy number in *E. coli*)	Urban & Vogel, 2007
pACYC184	carrying replication origin p15A	Chang & Cohen, 1978
pPHU231	broad-host-range expression vector	Hübner et al., 1991
pRK2013	helper plasmid for mobilization of non-self-transmissible plasmids	Figurski & Helinski, 1979
pLau44	pUC18 derivate carrying *tet* operator (*tetO*) cassettes	Lau et al., 2003
pART	artificial mini-replicon based on a pSymA derived *repA2B2C2* cassette	Frage et al., 2016
pK18mob2-repABC$_{pSymA}$	pK18mob2 carrying a pSymA derived *repA2B2C2* cassette	This work
p42b_p	pK18mob2 carrying *repABC*$_{p42b}$ based *oriV* (*R. etli*)	This work

p42d_p	pK18mob2 carrying *repABC*$_{p42d}$ based *oriV* (*R. etli*)	This work
p42e_s	pK18mob2 carrying a truncated *repABC* region derived from p42e (*R. etli*)	This work
p42f1_p	pK18mob2 carrying *repABC*$_{p42f1}$ based *oriV* (*R. etli*)	This work
p42f2_p	pK18mob2 carrying *repABC*$_{p42f2}$ based *oriV* (*R. etli*)	This work
pMla_s	pK18mob2 carrying a truncated *repABC* region derived from pMla (*M. loti*)	This work
pMlb_p	pK18mob2 carrying *repABC*$_{pMlb}$ based *oriV* (*M. loti*)	This work
pSMED01_p	pK18mob2 carrying *repABC*$_{pSMED01}$ based *oriV* (*S. medicae*)	This work
pSMED02_p	pK18mob2 carrying *repABC*$_{pSMED02}$ based *oriV* (*S. medicae*)	This work
pNGR234b_p	pK18mob2 carrying *repABC*$_{pNGR234b}$ based *oriV* (*S. fredii*)	This work
pTi_p	pK18mob2 carrying *repABC*$_{pTi}$ based *oriV* (*A. tumefaciens*).	This work
pMW1	integrative pK18mob2 derivative conferring Km^r	This work
pMW28	integrative pK18mob2 derivative carrying P*min2-eGFP*	This work
pJD120	pK18mobsacB derivative, integration of *tetR-YFP-mChr-parB*$_{Yp}$ under control of P*tauAB*	This work
pJD211	pK18mob2 derivative carrying P*min2-loxR-aacC1*	This work
pPHU231-eGFP	pPHU231 derivative, carrying a constitutive *eGFP* cassette	This work
pABCa-eGFP	pABCa carrying a constitutive *eGFP* cassette	This work
pABCa-eGFP-expR	pABCa carrying constitutively expressed *eGFP-expR*	This work
pABCa-tetO	pABCa carrying a *tetO* array	This work
pABCb-eGFP	pABCb carrying a constitutive *eGFP* cassette	This work
pABCb-eGFP-expR	pABCb carrying constitutively expressed *eGFP-expR*	This work
pABCc-eGFP-expR	pABCc carrying constitutively expressed *eGFP-expR*	This work
pABCc-parS$_{Yp}$	pABCc carrying *parS*$_{Yp}$	This work
pABCc-mob-eGFP	pABCc-mob carrying a constitutive *eGFP* cassette	This work
pABC1-eGFP	pABC1 carrying a constitutive *eGFP* cassette	This work
pABC1-eGFP-expR	pABC1 carrying constitutively expressed *eGFP-expR*	This work
pABC1-eGFPf	pABC1 carrying a forwardly integrated SD-*eGFP* fragment	This work
pABC1-eGFPr	pABC1 carrying a reversely integrated SD-*eGFP* fragment	This work
pABC1-T2-eGFPf	pABC1-T2 carrying a forwardly integrated SD-*eGFP* fragment	This work
pABC1-T2-eGFPr	pABC1-T2 carrying a reversely integrated SD-*eGFP* fragment	This work
pABC1-T3-eGFPf	pABC1-T3 carrying a forwardly integrated SD-*eGFP* fragment	This work
pABC1-T3-eGFPr	pABC1-T3 carrying a reversely integrated SD-*eGFP* fragment	This work
pABC2-eGFP	pABC2 carrying a constitutive *eGFP* cassette	This work
pABC2-eGFP-expR	pABC2 carrying constitutively expressed *eGFP-expR*	This work
pABC3-eGFP	pABC3 carrying a constitutive *eGFP* cassette	This work
pABC4-eGFP	pABC4 carrying a constitutive *eGFP* cassette	This work
pABC4-eGFP-expR	pABC4 carrying constitutively expressed *eGFP-expR*	This work
pABCb-lacI	pABCb carrying P*min2-lacI* for constitutive LacI expression	This work
pMlb_lacO	pK18mob2 carrying *repABC*$_{pMlb}$ bearing a *lacO* cassette at the	This work

	predicted transcription start site	
pIso	ABC cloning vector integrating upstream of the *exo* gene cluster on pSymB	This work
pIso-exo	isolated ABC cloning vector carrying the *exo* gene cluster (replicative upon IPTG supplementation)	This work
pABC library plasmids carrying module parts		
pLoriVSm	pK18mob2 derivative, unloaded library vector for oriVSm parts	This work
pLsynTer-MCS	pK18mob2 derivative, unloaded library vector for synTer-MCS parts	This work
pLoriT	pK18mob2 derivative, unloaded library vector for oriT parts	This work
pLAR	pK18mob2 derivative, unloaded library vector for AR parts	This work
pLoriV-42b	pLoriVSm derivative providing the $oriV_{p42b}$ region (*R. etli*)	This work
pLoriV-42d	pLoriVSm derivative providing the $oriV_{p42d}$ region *(R. etli)*	This work
pLoriV-42e	pLoriVSm derivative providing the $oriV_{p42e}$ region (*R. etli*)	This work
pLoriV-42f1	pLoriVSm derivative providing the *repA1B1C1* based $oriV_{p42f}$ region (*R. etli*)	This work
pLoriV-42f2	pLoriVSm derivative providing the *repA2B2C2* based $oriV_{p42f}$ region (*R. etli*)	This work
pLoriV-Mla	pLoriVSm derivative providing the $oriV_{pMla}$ region (*M. loti*)	This work
pLoriV-Mlb	pLoriVSm derivative providing the $oriV_{pMlb}$ region (*M. loti*)	This work
pLoriV-Ti	pLoriVSm derivative providing the $oriV_{pTi}$ region (*A. tumefaciens*)	This work
pLsynTer-1	pLsynTer-MCS derivative providing *synTer1*-MCS	This work
pLsynTer-2	pLsynTer-MCS derivative providing *synTer2*-MCS	This work
pLsynTer-3	pLsynTer-MCS derivative providing *synTer3*-MCS	This work
pLsynTer-1-lox	pLsynTer-MCS derivative providing *synTer1*-MCS-*loxL*	This work
pLsynTer-2-lox	pLsynTer-MCS derivative providing *synTer2*-MCS-*loxL*	This work
pLsynTer-3-lox	pLsynTer-MCS derivative providing *synTer3*-MCS-*loxL*	This work
pLoriT-1	pLoriT derivative providing a *mob* site	This work
pLoriT-1-lox	pLoriT derivative providing *mob-loxR*	This work
pLoriT-1-loxT	pLoriT derivative providing *mob-loxR-rrnBT2*	This work
pLAR-Km	pLAR derivative providing a kanamycin resistance cassette	This work
pLAR-Spec	pLAR derivative providing a spectinomycin resistance cassette	This work
pLAR-Tc	pLAR derivative providing a tetracyclin resistance cassette	This work
pLAR-Hyg	pLAR derivative providing a hygromycin resistance cassette	This work
pLAR-Tmp	pLAR derivative providing a trimethoprim resistance cassette	This work
pLAR-Gm	pLAR derivative providing a gentamicin resistance cassette	This work

Table 2. pABCs generated in this study.

		Modules						
	Plasmid	**oriVSm** (*oriV* origin)	**synTer-MCS** (relevant fragment)	**oriVEc** (*oriV* origin)	**oriT** (relevant fragment)	**AR** (resistance)	**Size** (kb)	**Design type**
No linker	**pABCa**	pMIb	*synTer1*	p15A	-	Gm	6.7	A
	pABCb	p42b	*synTer2*	pMB1	-	Sp	6.5	A
	pABCc	p42d	*synTer3*	pSC101*	-	Km	11.5	A
	pABCc-mob	p42d	*synTer3*	pSC101*	*mob*	Km	12.0	B
Module library based construction	**pABC1**	p42b	*synTer1*	pMB1	-	Sp	6.7	A
	pABC1-loxA	p42b	*synTer1-loxL*	pMB1	*mob-loxR*	Sp	9.5	C1
	pABC1-loxB	p42b	*synTer1-loxL*	pMB1	*mob-loxR-rrnBT2*	Sp	9.6	C2
	pABC1-sT2	p42b	*synTer2*	pMB1	-	Sp	6.7	A
	pABC1-sT3	p42b	*synTer3*	pMB1	-	Sp	6.7	A
	pABC2	pMIa	*synTer2*	p15A	-	Gm	7.2	A
	pABC3	pTi	*synTer3*	pSC101*	-	Hyg	8.6	A
	pABC4	p42e	*synTer1*	pMB1	-	Sp	7.0	A
	pABC5	p42f1	*synTer2*	pMB1	-	Sp	7.2	A
	pABC6	p42f2	*synTer2*	pMB1	-	Sp	7.3	A

pABCa-c are based on building blocks lacking the module-specific linker sequences and a transcription terminator downstream of the MCS.

Supplementary Material

A family of single copy *repABC*-type shuttle vectors stably maintained in the alpha-proteobacterium *Sinorhizobium meliloti*

Johannes Döhlemann[#], Marcel Wagner[#], Carina Happel, Patrick Sobetzko, and Anke Becker*

LOEWE Center for Synthetic Microbiology and Faculty of Biology, Philipps-Universität Marburg, Marburg, Germany

#These authors contributed equally.

*Correspondence: anke.becker@synmikro.uni-marburg.de

Supplementary Tables

Table S1. Examination of sequence identity of pABC module parts.

Part origin	pLoriV -42b	pLoriV -42d	pLoriV -42e	pLoriV -42f1	pLoriV -42f2	pLoriV -Mla	pLoriV -Mlb	pLoriV -Ti	pLsyn-Ter-1	pLsyn-Ter-2	pLsyn-Ter-3	pMB1	p15A	pSC101	pLoriT -1	pLAR -Km	pLAR -Gm	pLAR -Spec	pLAR -Tc	pLAR -Hyg	pLAR -Tmp
pLoriV-42b		0.76	0.78	0.82	0.82	0.76	0.79	0.73	0.53	0.53	0.53	0.55	0.54	0.54	0.54	0.55	0.54	0.54	0.57	0.56	0.53
pLoriV-42d	0.77		0.74	0.72	0.77	0.77	0.8	0.79	0.55	0.56	0.56	0.53	0.54	0.56	0.52	0.57	0.55	0.55	0.57	0.55	0.53
pLoriV-42e	0.78	0.75		0.75	0.8	0.79	0.8	0.71	0.52	0.54	0.52	0.53	0.55	0.55	0.53	0.58	0.57	0.55	0.56	0.55	0.52
pLoriV-42f1	0.82	0.63	0.75		0.83	0.73	0.79	0.69	0.56	0.54	0.55	0.53	0.53	0.56	0.54	0.57	0.56	0.57	0.56	0.57	0.53
pLoriV-42f2	0.82	0.77	0.8	0.83		0.79	0.81	0.71	0.54	0.53	0.55	0.54	0.55	0.53	0.54	0.56	0.57	0.55	0.55	0.56	0.53
pLoriV-Mla	0.76	0.77	0.79	0.73	0.79		0.87	0.75	0.53	0.54	0.54	0.52	0.54	0.53	0.57	0.56	0.52	0.52	0.56	0.57	0.56
pLoriV-Mlb	0.79	0.8	0.8	0.79	0.81	0.87		0.74	0.53	0.52	0.52	0.55	0.56	0.56	0.54	0.55	0.55	0.53	0.54	0.56	0.55
pLoriV-Ti	0.73	0.79	0.71	0.72	0.71	0.75	0.74		0.53	0.52	0.55	0.52	0.56	0.55	0.53	0.54	0.52	0.54	0.56	0.55	0.5
pLsynTer-1	0.53	0.55	0.52	0.56	0.54	0.53	0.53	0.53		0.8	0.8	0.51	0.52	0.51	0.53	0.53	0.51	0.52	0.52	0.52	0.5
pLsynTer-2	0.53	0.53	0.54	0.54	0.53	0.54	0.52	0.52	0.8		0.8	0.5	0.51	0.54	0.55	0.54	0.5	0.52	0.53	0.53	0.51
pLsynTer-3	0.53	0.56	0.52	0.55	0.55	0.54	0.52	0.55	0.83	0.77		0.49	0.52	0.48	0.56	0.53	0.53	0.51	0.53	0.51	0.51
$oriV_{pMB1}$	0.55	0.53	0.53	0.53	0.54	0.52	0.55	0.52	0.51	0.5	0.49		0.82	0.55	0.51	0.53	0.51	0.55	0.52	0.52	0.54
$oriV_{p15A}$	0.55	0.54	0.55	0.55	0.58	0.54	0.56	0.55	0.52	0.51	0.52	0.83		0.55	0.54	0.53	0.51	0.52	0.53	0.52	0.52
$oriV_{pSC101*}$	0.54	0.58	0.54	0.56	0.53	0.53	0.56	0.55	0.51	0.5	0.48	0.55	0.52		0.49	0.55	0.56	0.53	0.56	0.55	0.56
pLoriT-1	0.54	0.52	0.53	0.54	0.54	0.57	0.54	0.53	0.53	0.55	0.56	0.51	0.54	0.49		0.51	0.49	0.53	0.52	0.54	0.51
pLAR-Km	0.55	0.57	0.58	0.57	0.56	0.56	0.55	0.54	0.53	0.54	0.53	0.53	0.53	0.55	0.51		0.74	0.65	0.72	0.66	0.66
pLAR-Gm	0.55	0.55	0.57	0.56	0.57	0.52	0.55	0.52	0.51	0.5	0.53	0.51	0.51	0.55	0.49	0.74		0.73	0.62	0.72	0.72
pLAR-Spec	0.54	0.55	0.55	0.57	0.55	0.52	0.53	0.54	0.52	0.52	0.51	0.55	0.52	0.53	0.53	0.65	0.73		0.68	0.65	0.71
pLAR-Tc	0.57	0.57	0.56	0.56	0.55	0.56	0.54	0.56	0.51	0.51	0.53	0.52	0.51	0.56	0.52	0.72	0.62	0.68		0.71	0.69
pLAR-Hyg	0.56	0.55	0.55	0.57	0.56	0.57	0.56	0.55	0.52	0.53	0.51	0.52	0.52	0.55	0.54	0.66	0.72	0.65	0.72		0.7
pLAR-Tmp	0.53	0.53	0.52	0.53	0.53	0.56	0.55	0.5	0.5	0.51	0.51	0.54	0.52	0.56	0.51	0.66	0.72	0.71	0.69	0.7	

>70% Seq. Id.
>80% Seq. Id.
>85% Seq. Id.

pABCs have been designed on the premise of low sequence identity in order to avoid homologous recombination through regions sharing sequence similarities between distinct pABCs. Sequence identity analysis of 200 bp sliding windows aligning parts of all library plasmids (and *E. coli oriV*s pMB1, p15A and pSC101*) were performed. Analysis revealed 12.5%, 4.3% and 0.5% of 441 possible combinations exhibiting a maximum sequence identity of 70-80%, 80-85% and >85%, respectively. Since homologous recombination in *S. meliloti* commonly occurs within much larger regions, pABCs composed of distinct parts are not supposed to fuse with an appreciable frequency. Seq. Id.: sequence identity. Maximal values are indicated. pMB1: $oriV_{pMB1}$. p15A: $oriV_{p15A}$. pSC101: $oriV_{pSC101*}$.

Table S2 (1/7). Primers and synthetic DNA fragments used in this study.

Number/ Name	Sequence (5‘-3‘)	Used for
BO1	AACTTCTCGTGGAACTCTATCGTAACTGAAGTCTATTGAAGGAACACTGTATCTCGGTCA	pABC assembly
BO2	GTTATCAGAAGAGTGTAAGCGGGTTGACTATACTACTGTTTCAGACTGGCGTAATCACTC	pABC assembly
BO3	GGCAATCACTGAGTAAGACTATTCTGTCCAGACCTTTTCTCCGACGAATAGAGTAACAGA	pABC assembly
BO3a	GGCAATCACTGAGTAAGACTATTCTGTCCATGTGACGATAAGTTCCCTACTGACAGAATC	pABC assembly
BO3b	GGTTACAGTGTCACCCGTTATCAATAGAGGGACCTTTTCTCCGACGAATAGAGTAACAGA	pABC assembly
BO4	GATTGGTCAGGCACAACTTTCAGATAAGACGAAACTGTCACTGGTCCGTTGATACAATCC	pABC assembly
Uni	AGCGGATAACAATTTCACACAGGA	Binds specifically to the pK18mob2 backbone; used for CFGE and colony PCR
Rev	AAGGCGATTAAGTTGGGTAACG	
2	CATCTGCAGTTCAGGAGGAATAACTATGGTGAGCAAGGGCGAGGAG	PCR amplification of promoterless *eGFP*
3	GTACAAGCTTTTACTTGTACAGCTCGTCCATG	
6	CGCGGATCCCCTGACGATGCGTGGAGACC	PCR amplification of P*aacC1*
7	GCGCATATGCGTAACGCGCTTGCTGCTTG	
15	CGCCATATGATAACTTCGTATAGCATACATTATACGAAGTTATGTCGACCGC	Generation of *loxP*
16	GCGGTCGACATAACTTCGTATAATGTATGCTATACGAAGTTATCATATGGCG	
21	CGCACCGGCCTCAATATTGC	PCR amplification of a homologous region upstream of the *exo* gene cluster
22	CAACCTCTGGGGATGATTTC	
48	GCGGTACCTGCCCGTGCCGATTGTGTCC	PCR amplification of *repA2B2C2* based *ori*V_{pSymA}
50	CGGGATCCAGAACCATCCGGTGCAGTCG	
57	CGGTCGACCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGT	Generation of *rrnBT1*
58	ACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGGTCGACCG	
59	TTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCTGCAGCG	
60	CGCTGCAGATTTGTCCTACTCAGGAGAGCGTTCACCGACAAACAACAGATAAA	
61	CGCCATATGCATCAAATTAAGCAGAAGGCCATCCTGAC	Generation of *rrnBT2*
62	GTCAGGATGGCCTTCTGCTTAATTTGATGCATATGGCG	
63	GGATGGCCTTTTTGCGTTTCTACAAACTCTTGTCGACCG	
64	CGGTCGACAAGAGTTTGTAGAAACGCAAAAAGGCCATCC	
78	GGATCGAGGCGTTGAACTTC	qPCR (primer bind specifically to the chromosome)
79	GATAAGCCCGACGATGAGTG	
80	CAGCATTGGCAGCATTGATC	qPCR (primer bind specifically to pSymA)

Table S2 (2/7). Primers and synthetic DNA fragments used in this study.

Number/ Name	Sequence (5‘-3‘)	Used for
80	CAGCATTGGCAGCATTGATC	qPCR (binds specifically to pSymA)
81	GTTCGAGCCGAACGATCATC	
90	GCGGAATTCAACAGCCAGATGCTCGATACC	PCR amplification of a homologous region in *exoP*
91	GCGGGATCCGACGATGAGCTGGCTGTTGAG	
92	CGCGAATTCCATCAAAAGCGCTGTAGCTGC	PCR amplification of a homologous region *exoI*
93	GCGGGATCCCAACTCGGCTTTCGGTTCACC	
94	CGCGAATTCAATGCGGCGATGAAGTCATGC	PCR amplification of a homologous region spanning *exsA*
95	GCGGGATCCTTCGTGCTCACCTCGTTCTTC	
125	GCGGTCGACGTTATGGAGCAGCAACGATGTT	PCR amplification of *aacC1*
126	GCGCTGCAGAGTACTTTAGGTGGCGGTACTTGGGTCG	
127	CGCGTCGACGGACTCAATACACCATGAGGG	PCR amplification of *aadA1*
128	GCGCTGCAGAGTACTTTATTTGCCGACTACCTTGGTG	Binds to *aadA1*; used for PCRs with pIso-exo as a template
131	CGCGGATCCCTGACGCGTACAGGAAACAC	PCR amplification of *rpoC-thrA* transcription terminator
132	GCGCATATGCGGGCTCTTTCCCTAAACTCC	
137	GTATTGGTCGCGCTCTTCC	PCR amplification of *repABC* based $oriV_{pTi}$
138	GTTGATCACCGACGTCGAC	
139	CGCCATCAAACCAAGCGAG	PCR amplification of *repABC* based $oriV_{pMLa}$
140	CTGTCATCAACAAGCCGCG	
141	GGCGTTCGTCACTCGTTTG	PCR amplification of *repABC* based $oriV_{pMlb}$
142	CACCGACTGCCATGTGATC	
145	GGCCCGAGATCAATTCTCC	PCR amplification of *repABC* based $oriV_{p42e}$
146	CGACTGGCGCAAAGCTTTC	
149	CCAGAGGCCAGAGAGAATG	PCR amplification of *repABC* based $oriV_{pNGR234b}$
150	CCTCGTCCGTTTCCTTCAG	
151	CTGAGTGCGTCTTGAAGCC	PCR amplification of *repABC* based $oriV_{pSMED01}$
152	CACGGTGTTGTAGAGCAGG	
153	CTAGCAAGGACGAGACACG	PCR amplification of *repABC* based $oriV_{pSMED02}$
154	CCTGAACCGGTCATTGTCG	

Table S2 (3/7). Primers and synthetic DNA fragments used in this study.

Number/ Name	Sequence (5'-3')	Used for
187	TATGTACCGTTCGTATAGCATACATTATACGAAGTTATG	Generation of *loxL*
188	TCGACATAACTTCGTATAATGTATGCTATACGAACGGTACA	
189	TATGATAACTTCGTATAGCATACATTATACGAACGGTAG	Generation of *loxR*
190	TATGTACCGTTCGTATAGCATACATTATACGAACGGTAG	
191	TATGTACCGTTCGTATAGCATACATTATACGAACGGTAG	Generation of *loxLR*
192	TCGACTACCGTTCGTATAATGTATGCTATACGAACGGTACA	
211	CTTTCTGAAGATTGCCGATC	PCR amplification of *repABC* based $oriV_{p42b}$
212	CGTCAAAAAGGCTTTTCACG	
215	GAGACGGAAATACGTCTACC	PCR amplification of *repABC* based $oriV_{p42d}$
216	GACGGTACCTATCTCTACAC	
217	GATGACGGAAAATGGCTTTC	PCR amplification of *repA1B1C1* based $oriV_{p42f}$
218	CATCACAGTCGGCCTTTATG	
219	GTGATCATTCACTCACCTCG	PCR amplification of *repA2B2C2* based $oriV_{p42f}$
220	GGTTGCTAATCGGGACTAAG	
236	CGCGAATTCTGGAGACCGGGTCAAGCTG	PCR amplification of *tetR*
244	CAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAAGTTAAGTATAGTAGAGGAGAAAACA	Generation of P*min1*
245	TATGTTTTCTCCTCTACTATACTTAACTTACAGTTTACGAACCGAACAGGCTTATGTCAACTGCATG	
247	CGGGATCCTTAGGTGGCGGTACTTGGGTC	PCR amplification of *aacC1*
249	CGGGATCCTTATTTGCCGACTACCTTGGTGATCT	PCR amplification of *aadA1*
250	CGGGTACCAATGCGGCGATGAAGTCATG	PCR amplification of a homologous region comprising *exsA*
251	CGGAATTCTTCGTGCTCACCTCGTTCTT	
252	CTCGCAGAGCAGGATTCCCGTT	qPCR, binds specificly to the *mob* site
253	GGCAGGATAGGTGAAGTAGGCCCA	
255	CGGGATCCCTATTCCTTTGCCCTCGGACG	PCR amplification of *hph*
259	CGGGATCCTTAGGCCACGCGTTCAAGC	PCR amplification of *tmpR*
261	CGGGATCCTCAGAAGAACTCGTCAAGAA	PCR amplification of *nptII*
266	GCGGAATTCGAAGGTCGAGCGCAACTTG	PCR amplification of a homologous region comprising the upstream region of *tauA*
267	GCGGTCGACCGTTACCCTCTTTGGTTATGTC	
269	GCGGCATGCGTTGAGGATGTTCACTTCC	PCR amplification of a homologous region comprising P*tauAB*

Table S2 (4/7). Primers and synthetic DNA fragments used in this study.

Number/ Name	Sequence (5‘-3‘)	Used for
270	GCGGTCGACTAGCAGGAGGAATTCACCAT	PCR amplification of *tetR-YFP*
271	GCGGCTAGCTTATCTAGACTTGTACAGCTCG	
272	GCGGCTAGCGGATCAATCCGAGGAAAAAC	PCR amplification of *mChr-parB*$_{Yp}$
273	CGCGGCATGCAGTACTTTACTCACCTGATTCTGGAAG	
284	CTGTTGACAATTAATCATCGAACTAGTTAACTAGTACGCAAGTAGACGCGTTACGTAGCAGAGGAGAAAACA	PCR amplification of P*min2* based antibiotic resistance/*eGFP* cassettes; Generation of P*min2*
285	TATGTTTTCTCCTCTGCTACGTAACGCGTCTACTTGCGTACTAGTTAACTAGTTCGATGATTAATTGTCAACAGCATG	Generation of P*min2*
286	CATAACGGTTCCGGCAAATATACTGAAATAGGTGTTGACATTATTCCATCGAACTAGTTAAC	Generation of P*min3*
287	TAACTAGTTCGATGGAATAATGTCAACACCTATTTCAGTATATTTGCCGGAACCGTTATGCATG	
288	TAGTACGAAAGTTCACATGAATTCGATGTTGAGTTCAGAGGAGAAAACA	
289	TATGTTTTCTCCTCTGAACTCAACATCGAATTCATGTGAACTTTCGTACTAGT	
290	CGCCATATGTTACGCAGCAGCAACGATGTTAC	PCR amplification of *aacC1*
291	CGCCATATGAGGGAAGCGGTGATCGCC	PCR amplification of *aadA1*
292	CGCCATATGGACCAACACAACAATGGAGTCAGT	PCR amplification of *tmpR*
301	TCGACCATCAAATTAAGCAGAAGGCCATCCTGACGGATGG	Generation of *rrnBT2*
302	CCATCCGTCAGGATGGCCTTCTGCTTAATTTGATGG	
326	CGCCATATGGTGAGCAAGGGCGAGGAG	PCR amplification of *eGFP*
327	CGCGGATCCTTACTTGTACAGCTCGTCCATGC	PCR amplification of *SD-eGFP*
330	GCGAGTACTACGGTCGAGCCCGCCATTGTTC	PCR amplification of a homologous region comprising P*tauAB*
331	CTCGCGGACGTGCTCATAG	PCR amplification of the pK18mob2 derived *mob* site
332	GCTTTTCCGCTGCATAACCC	
333	CTACATTTGAAGAGATAAATTGCA	PCR amplification of *oriV*$_{p15A}$
334	TAGCGGAGTGTATACTGG	
335	CGCGTTGCTGGCGTTTTT	PCR amplification of *oriV*$_{pMB1}$
336	AGATCAAAGGATCTTCTTGAGATCCTT	
338	CCTAGGGTACGGGTTTTG	PCR amplification of *oriV*$_{pSC101}$
340	GCCTTTTTGCGTTTCTACAAACTCTTCCCGGCGTTCGTCACTCGTTTGGCC	LCR assembly of pABCa
341	GATTATCTAGATCACATGGCAGTCGGTGCTGTTGACAATTAATCATCGAACTAGTTAACTAGTA	
344	CTAGATTTCAGTGCAATTTATCTCTTCAAATGTAGGGGCATATGAGTACTGCGGCCG	
345	AAGTACCGCCACCTAAGGATCCCGTAGCGGAGTGTATACTGGCTTACTATGTT	

Table S2 (5/7). Primers and synthetic DNA fragments used in this study.

Number/ Name	Sequence (5‘-3‘)	Used for
346	GCCTTTTTGCGTTTCTACAAACTCTTCCCCTTTCTGAAGATTGCCGATCCCGATAT	LCR assembly of pABCb
347	GGAGTTGACGTGAAAAGCCTTTTTGACGCTGTTGACAATTAATCATCGAACTAGTTAACTAGTA	
350	GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTGGGCATATGAGTACTGCGGCCG	
351	AAGGTAGTCGGCAAATAAGGATCCCGCGCGTTGCTGGCGTTTTTCCAT	
352	GCCTTTTTGCGTTTCTACAAACTCTTCCCGAGACGGAAATACGTCTACCAGCATGG	LCR assembly of pABCc(-mob)
353	AGATCGGTTGTGTAGAGATAGGTACCGTCCTGTTGACAATTAATCATCGAACTAGTTAACTAGTA	
354	CGCGTGGCCTAAGGATCCCGCTCGCGGACGTGCTCATAGTCC	LCR assembly of pABCc-mob
366	CGCGCATGCTACGTAGCAGAGGAGAAAACATATG	PCR amplification of *SD-eGFP*
367	CGCCATATGAAATCTAACAATGCGCTCATCGTCATC	PCR amplification of *tetA*
368	CGCCATATGAAAAAGCCTGAACTCACCGCG	PCR amplification of *hph*
369	CGCCATATGATTGAACAAGATGGATTGCA	PCR amplification of *nptII*
380	GACAGTAAGACGGGTAAGC	PCR amplification of $oriV_{pSC101*}$
381	GCAGGGTTATGCAGCGGAAAAGCCCTAGGGTACGGGTTTTGCTGCCC	LCR assembly of pABCc-mob
382	CATCAACAGGCTTACCCGTCTTACTGTCGGGCATATGAGTACTGCGGCCG	LCR assembly of pABCc(mob)
383	CGCGTGGCCTAAGGATCCCGCCTAGGGTACGGGTTTTGCTGCCC	
388	GATCCTGTTGACAATTAATCATCGAACTAGTTAACTAGTACGCAAGTAGACGCGTTACGTCA	Generation of P*min2*
389	TATGACGTAACGCGTCTACTTGCGTACTAGTTAACTAGTTCGATGATTAATTGTCAACAG	
436	CGCGGTACCTCAGGTCGAGGTGGCCC	PCR amplification of *tetA*
448	CGCGGCGGCCGCTGTTTTTCACCACGCCAATT	PCR amplification of $parS_{Yp}$
449	CGCGGCGGCCGCGAGGTTGAAAAGCGTGGTGA	
462	GAAACTGTCACTGGTCCGTTGATACAATCCAGTACTAACTTCTCGTGGAACTCTATCGTAACTGAAA	Generation of the oriVSm linker fragment
463	AGCTTTTCAGTTACGATAGAGTTCCACGAGAAGTTAGTACTGGATTGTATCAACGGACCAGTGACAGTTTC	
470	GAGTGAGACCCTCCCCTTC	PCR amplification of $repABC_{p42e}$
471	GCGAAGAACTTGGCAAGGC	PCR amplification of $repABC_{pMla}$
494	GAAACTGTCACTGGTCCGT	Standardized primers for PCR amplification of oriVSm module parts from library plasmids
495	TTCAGTTACGATAGAGTTCCACG	
496	GTCTATTGAAGGAACACTGTATCTCG	Standardized primers for PCR amplification of synTer-MCS module parts from library plasmids
497	TAGTCAACCCGCTTACACTC	
500	TGTGACGATAAGTTCCCTACTG	Standardized primers for PCR amplification of oriT module parts from library plasmids
501	CCTCTATTGATAACGGGTGACA	
502	GACCTTTTCTCCGACGAATAGA	Standardized primers for PCR amplification of AR module parts from library plasmids

Table S2 (6/7). Primers and synthetic DNA fragments used in this study.

Number/ Name	Sequence (5‘-3‘)	Used for
503	GTCTTATCTGAAAGTTGTGCCTG	Standardized primers for PCR amplification of AR module parts from library plasmids
544	TACTACTGTTTCAGACTGGCGTAATCACTCAGTCTACATTTGAAGAGATAAATTGCACTG	PCR amplification of oriVEc module part *oriV*$_{p15A}$ from e.g. pSC101
545	TGGACAGAATAGTCTTACTCAGTGATTGCCAGTTAGCGGAGTGTATACTGGC	
546	TACTACTGTTTCAGACTGGCGTAATCACTCAGTAGATCAAAGGATCTTCTTGAGATCCTTTTTTTC	PCR amplification of oriVEc module part *oriV*$_{pMB1}$ from e.g. pK18mob2
547	TGGACAGAATAGTCTTACTCAGTGATTGCCAGTCGCGTTGCTGGCGTTTTTC	
573	GACGAGTTCTTCTGAGCGGG	PCR amplification of *sacB-mob*
574	ACTTACCGTTCGTATAGCATACATTATACGAAGTTATCA	Generation of *loxLE*
575	TATGATAACTTCGTATAATGTATGCTATACGAACGGTAAGT	
576	GCGCATATGGCTTTTCCGCTGCATAACCC	PCR amplification of *sacB-mob*
577	ACTCATATGATAACTTCGTATAGCATACATTATACGAACGGTAG	Generation of *loxRE-rrnBT2*
578	TCGACTACCGTTCGTATAATGTATGCTATACGAAGTTATCATATGAGT	
579	TCGACAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAATTTGATG	
580	CATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTG	
581	ACTCATATGATAACTTCGTATAGCATACATTATACGAACGGTA	Generation of *loxRE*
582	TACCGTTCGTATAATGTATGCTATACGAAGTTATCATATGAGT	
608	CGCTTAATTAAATTGAATTCTGATTGACACCATCGAATGGC	PCR amplification of P*lacI*
609	CGCGTTTAAACTCACTGCCCGCTTTCCAGT	PCR amplification of *lacI*
613	CGCCATATGAAACCAGTAACGTTATACGATGTCGCAG	
614	CGCCATATGTTTTCTCCTCAATTGACTCTCTTCCGGGCGC	PCR amplification of P*lacI*
629	AATTGTTATCCGCTCACAATTAAGCGCCATTAAGCTCCGATTC	Overlap-PCR for *repABC*$_{pMb}$*-lacO* (with primers 141+142)
630	AATTGTGAGCGGATAACAATTGTTGTGGGCATTGCTTACAAAAAG	
647	CGCGGTACCTGTTGACAATTAATCATCGAACTAGTTAAC	PCR amplification of *eGFP*
648	CGCGAATTCATTGCGGCCGCTTACTTGTACAGCTCGTCCA	
649	CGCGCGGCCGCCTGGAAAGCGGGCAGTGAG	PCR amplification of *eGFP-expR*
651	CGCTTAATTAATGTTGACAATTAATCATCGAACTAGT	
652	CGCAGTACTTCAGGAGATCAGTCCCAGGC	
670	TACTACTGTTTCAGACTGGCGTAATCACTCAGTCCTAGGGTACGGGTTTTGC	PCR amplification of oriVEc module part *oriV*$_{pSC101*}$ from e.g. pXG-10
671	TGGACAGAATAGTCTTACTCAGTGATTGCCAGTGACAGTAAGACGGGTAAGCCT	
680	CGACGGCATTCTCTTCGTGGT	Binds to *exoP*, used for PCR with pIso-exo as a template
717	CGCTCTAGATGATTGACACCATCGAATGGCG	PCR amplification of *lacI*
718	CGCTCTAGATCACTGCCCGCTTTCCAGT	

Table S2 (7/7). Primers and synthetic DNA fragments used in this study.

Number/ Name	Sequence (5‘-3‘)	Used for
module_linkers	GGGTACTACTGTTTCAGACTGGCGTAATCACTCAGTACTGGCAATCACTGAGTAAGACTATTCTGTCCAGG ATCCCGGGGTCTATTGAAGGAACACTGTATCTCGGTCAAGTACTAAGAGTTTGTAGAAACGCAAAAAGGC CATCCGTCAGGATGGCCTTCTGCTTAATTTGATGGTTATCAGAAGAGTGTAAGCGGGTTGACTATCTAGAT GTGACGATAAGTTCCCTACTGACAGAATCAGTACTGGTTACAGTGTCACCCGTTATCAATAGAGGCTGCAG ATCCCGGGGACCTTTTCTCCGACGAATAGAGTAACAGAAGTACTGATTGGTCAGGCACAACTTTCAGATAA GACAAGCTT	Release of individual pABC module linker sequences
synTer1_pre	GGGCATATGAGTACTGCGGCCGCTTAATTAAGTTTAAACCCCGACGACGGTCGGCTTCCTCGCCGAGAGT TCGGGCTCCGGCGGCGCGGAGGTCTTCGGCGCGATCGGCTACGGGATCGCGGATTTGTCCTACTCAGG AGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTG ATGCGCCGGTTATCTCGGGGCTCCCGGAGACCCGCGGATTCGAACTCGGAAGAACCGTTCGGACTCCGC GACCGTTAACCGGCCGCTCGCCCTCGTTTACGAGGACGCGGACGCTCCGAGTTCCATCAAATTAAGCAG AAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTGAAGACGGTTCGCGTTCGCTTCCGG AAGGGGCCGACCGGCCCGACGGCCGGATTCGCCGCCGGCGAGACCGAGGCGAACCTCTTCTCGCTTTC CGCCGAGTCGAC	Construction of the synTer-MCS module (*synTer1*)
synTer2_pre	GGGCATATGAGTACTGCGGCCGCTTAATTAAGTTTAAACCTCGACGAGCTTCTAAACGACGGTCTAATCGA GGAGCTCCGCGAGGGCGAGGCCGAGCCCGATTATCGGCGAGAAGCCGCCCGATTTGTCCTACTCAGGA GAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGA TGTCGTCTAGGCGCCGAACTCCTCCGCGAGGAGGAGCGCGGCGGAGGTCTCCTCCTAAGCCGGACTAA GGAGACCGTAATCCTCGGCCTCCCGGTCGTCGCCGGTCCGCTCCGGGGCCTCGCCATCAAATTAAGCAG AAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTGCTTCCGGGCGAGGGTCTACTACTC GTAGTCCGTCGCGAGTCCGAACGCGGTCGCGGTCTAGAGCGGTCGCTAAGAGAGGCTCTAAGCGCCGA GCCCCTGTCGAC	Construction of the synTer-MCS module (*synTer2*)
synTer3_pre	GGGCATATGAGTACTGCGGCCGCTTAATTAAGTTTAAACTCCTCGGGACGCTCGCCGAGCTCCTCTCCGA GGAGCGGCCCGCCGCCGGACGTCGATAACGGAGACCGCGCTCCCGTCGATCCATTTGTCCTACTCAGGA GAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGA TGCGTCGAGCCGGTCGGCCGCGTCGTCGATAACTAACTCCTCGGCGTCGTCGTCGAGACTAACCTCGAG GGCGACGGAGACGACGAGGGCCCGTCCGTCCTCGGCGACCGTCCCCCGCGTCGCATCAAATTAAGCAG AAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTCGTACGGGTCTAAGACCTTAACGACG CCCGCGACCTCTAAGAGGTCGGCGACGTAACGCCGGACCTCGGCCCGTACCGCTAACTCTAAGACGTAA GCGGGTCGAC	Construction of the synTer-MCS module (*synTer3*)

Underlined: *rrnBT1/2* transcription terminators

Table S3. Designation of pABC library parts according to the SEVA-SIB nomenclature.

Module	Position	Part origin (library plasmid)	*	Genbank accession number
AR	1	pLAR-Km	1	KY031715
		pLAR-Spec	2	KY031716
		pLAR-Tc	3	KY031717
		pLAR-Hyg	4	KY031718
		pLAR-Tmp	5	KY031719
		pLAR-Gm	6	KY031720
oriVEc	2	(pMB1)	1	KY031721
		(p15A)	2	KY031722
		(pSC101*)	3	KY031723
oriVSm	3	pLoriV-42b	1	KY031698
		pLoriV-42d	2	KY031699
		pLoriV-42e	3	KY031700
		pLoriV-42f1	4	KY031701
		pLoriV-42f2	5	KY031702
		pLoriV-Mla	6	KY031703
		pLoriV-Mlb	7	KY031704
		pLoriV-Ti	8	KY031705
synTer -MCS	4	pLsynTer-1	1	KY031706
		pLsynTer-2	2	KY031707
		pLsynTer-3	3	KY031708
		pLsynTer-1-lox	4	KY031709
		pLsynTer-2-lox	5	KY031710
		pLsynTer-3-lox	6	KY031711
oriT	5	pLoriT-1	A	KY031712
		pLoriT-1-lox	B	KY031713
		pLoriT-1-lox-T	C	KY031714

Modules are assigned to five positions, and parts receive individual identifiers (numbers or letters). *: Position designation. (pMB1): oriV derived from pK18mob2. (p15A): *oriV* derived from pACYC184. (pSC101*): *oriV* derived from pXG-10.

Table S4. Translation of pABC plasmids into the SEVA-SIB collection.

Position		**1**	**2**	**3**	**4 + 5**	
Trivial name	**pSEVA**	Resistance (AR)*	ORI (oriVEc)*	CARGO-LIKE (oriVSm)*	GADGET (synTer-MCS + oriT)*	**Genbank accession number**
pABCa	**pSEV6271**	Gm	p15A	pMIb	*synTer1*	KY031724
pABCb	**pSEV2112**	Sp	pMB101	p42b	*synTer2*	KY031725
pABCc	**pSEV1323**	Km	pSC101*	p42d	*synTer3*	KY031726
pABCc-mob	**pSEV1323A**	Km	pSC101*	p42d	*synTer3 + mob*	KY031727
pABC1	**pSEV2111**	Sp	pMB1	p42b	*synTer1*	KY031728
pABC1-loxA	**pSEV2114B**	Sp	pMB1	p42b	*synTer1-loxL + mob-loxR*	KY031729
pABC1-loxB	**pSEV2114C**	Sp	pMB1	p42b	*synTer1-loxL + mob-loxR-rrnBT2*	KY031730
pABC1-sT2	**pSEV2112**	Sp	pMB1	p42b	*synTer2*	KY031731
pABC1-sT3	**pSEV2113**	Sp	pMB1	p42b	*synTer3*	KY031732
pABC2	**pSEV6262**	Gm	p15A	pMIa	*synTer2*	KY031733
pABC3	**pSEV4383**	Hyg	pSC101*	pTi	*synTer3*	KY031734
pABC4	**pSEV2131**	Sp	pMB1	p42e	*synTer1*	KY031735
pABC5	**pSEV2142**	Sp	pMB1	p42f1	*synTer2*	KY031736
pABC6	**pSEV2152**	Sp	pMB1	p42f2	*synTer2*	KY031737

Dependent on the number of module parts, pSEV derivatives are composed of 4 to 5 positions. Position identifiers are defined in Table S3. Module AR is assigned to position 1 (Resistance). Module oriVEc is described in position 2 (ORI). Position 3 (CARGO-LIKE) indicates the oriVSm module part. Position 4 + 5 (GADGET) are combined and denote the synTer-MCS (numbers) and oriT (letters) parts. (AR)*: conferred antibiotic resistances. Gm: gentamicin; Sp: spectinomycin; Km: kanamycin; Hyg: hygromycin; (oriVEc)*: *E. coli oriV* origin is indicated. (oriVSm)*: *S. meliloti oriV* megaplasmid origin is indicated. (synTer-MCS + oriT)*: relevant inserts are indicated.

Table S5. qPCR indicates a unit copy number of plasmids carrying a *repABC*$_{pSymA}$ cassette.

Plasmid	[DNA]*	Ct (Chr)	StDev (Ct(Chr))	Ct (pSymA)	StDev (Ct(pSymA))	ΔCt	ΔCt ref	ΔΔCt	Normalized Copy Number	Ø Copy Number	StDev
(Rm1021)	5	24.52	0	24.28	0.23	-0.24	-0.24	0	1	-	-
(Rm1021)	1	24.55	0.08	25.56	0.06	1.01	1.01	0	1		
(Rm1021)	0.1	25.96	0.37	26.88	0.12	0.92	0.92	0	1		
pK18mob2-*repABC*$_{pSymA}$	5	19.71	0.49	20.32	0.36	0.61	-0.24	0.85	0.56	**0.91**	**0.31**
pK18mob2-*repABC*$_{pSymA}$	1	19.81	0.30	20.66	0.22	0.85	1.01	-0.16	1.12		
pK18mob2-*repABC*$_{pSymA}$	0.1	23.09	0.23	23.93	0.12	0.85	0.92	-0.07	1.05		
pART	5	25.19	0.26	25.86	0.04	0.66	-0.24	0.90	0.54	**0.65**	**0.32**
pART	1	26.99	0.13	27.99	0.07	0.99	1.01	-0.02	1.01		
pART	0.1	30.14	0	32.39	0.11	2.25	0.92	1.33	0.40		
Plasmid	**[DNA]***	**Ct (Chr)**	**StDev (Ct(Chr))**	**Ct (*mob*)**	**StDev (Ct(*mob*))**	**ΔCt**	**ΔCt ref**	**ΔΔCt**	**Normalized Copy Number**	**Ø Copy Number**	**StDev**
pK18mob2-*repABC*$_{pSymA}$	5	17.89	0.37	25.72	0.51	7.83	7.83	0	1	-	-
pK18mob2-*repABC*$_{pSymA}$	1	19.82	0.21	28.19	0.07	8.37	8.37	0	1		
pK18mob2-*repABC*$_{pSymA}$	0.1	22.80	0.17	32.19	0.18	9.39	9.39	0	1		
pPHU231	5	18.72	0.01	23.62	0.25	4.90	7.83	-2.93	7.62	**5.56**	**1.87**
pPHU231	1	20.34	0.28	26.36	0.26	6.01	8.37	-2.35	5.11		
pPHU231	0.1	23.10	0.13	30.50	0.39	7.40	9.39	-1.99	3.96		

Cell lysates of *S. meliloti* Rm1021 and *S. meliloti* Sma818 carrying pK18mob2-*repABC*$_{pSymA}$ and pART, respectively, were generated from cell cultures growing in the late exponentially phase (OD_{600} ~ 1). The copy number was determined from the ratio of a chromosomal sequence (Chr; primers 78 and 79) to a target sequence within *repABC*$_{pSymA}$ (pSymA, primers 80 and 81) and *S. meliloti* Rm1021 served as reference strain (grey background). Both pK18mob2/*repABC*$_{pSymA}$ and pART exhibited a similar copy number close to 1. This was expected since both plasmids are based on the same *repABC* cassette. Fluorescence microscopy examination further verified single copy replication of pART (Frage et al., 2016) and thus the caculated copy number of pK18mob2-*repABC*$_{pSymA}$. Since pK18mob2 and pPHU231 have the same *mob* site in common, pK18mob2-*repABC*$_{pSymA}$ was applied as reference replicon for copy number determination of pPHU231. Accordingly, the copy number was determined from the ratio of a chromosomal reference sequence (Chr; primers 78 and 79) to a target sequence within the *mob* site (*mob*, primers 252 and 253). pPHU231 was expected to exhibit a low copy number in the examined range (Hübner et al., 1991). StDev: standard deviation. R1/2: biological replicates. [DNA]*: DNA concentration was estimated from the optical density of *S. meliloti* cell cultures. ref: Reference.

Frage, B., Döhlemann, J., Robledo, M., Lucena, D., Sobetzko, P., Graumann, P. L., and Becker, A. (2016) Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. *Mol. Microbiol.100*, 808-823.

Hübner, P., Willison, J. C., Vignais, P. M., and Bickle, T. A. (1991) Expression of regulatory *nif* genes in *Rhodobacter capsulatus*. *J.Bacteriol. 173*, 2993–2999.

Table S6. Copy number determination of oriVSm library plasmids.

Plasmid	[DNA]	Ct (Chr)	StDev (Ct(Chr))	Ct (*mob*)	StDev (Ct(*mob*))	ΔCt	ΔCt ref	ΔΔCt	Normalized Copy Number	Ø Copy Number	StDev
pK18mob2-*repABC*$_{pSymA}$	1	20.36	0.09	22.86	0.41	2.50	2.50	0.00	1	-	-
pK18mob2-*repABC*$_{pSymA}$	0.1	23.31	0.27	26.20	0.35	2.89	2.89	0.00	1		
pLoriV-42b R1	1	20.42	0.17	23.30	0.26	2.89	2.50	0.39	0.8	1.0	0.2
pLoriV-42b R1	0.1	23.77	0.80	26.46	0.25	2.68	2.89	-0.20	1.2		
pLoriV-42b R2	1	20.72	0.09	23.00	0.17	2.28	2.50	-0.22	1.2		
pLoriV-42b R2	0.1	23.77	0.34	26.72	0.28	2.95	2.89	0.06	1.0		
pLoriV-42d R1	1	20.32	0.33	23.06	0.33	2.74	2.50	0.24	0.8	1.1	0.2
pLoriV-42d R1	0.1	24.17	0.11	26.89	0.22	2.72	2.89	-0.17	1.1		
pLoriV-42d R2	1	21.20	0.09	23.66	0.44	2.46	2.50	-0.04	1.0		
pLoriV-42d R2	0.1	24.91	0.34	27.49	0.34	2.58	2.89	-0.31	1.2		
pLoriV-42e R1	1	21.07	0.32	23.15	0.11	2.08	2.50	-0.42	1.3	1.2	0.4
pLoriV-42e R1	0.1	24.10	0.29	26.79	0.21	2.69	2.89	-0.19	1.1		
pLoriV-42e R2	1	21.26	0.22	24.03	0.27	2.77	2.50	0.27	0.8		
pLoriV-42e R2	0.1	25.01	0.08	27.15	0.16	2.14	2.89	-0.75	1.7		
Strain	**[DNA]**	**Ct (Chr)**	**StDev (Ct(Chr))**	**Ct (*mob*)**	**StDev (Ct(*mob*))**	**ΔCt**	**ΔCt ref**	**ΔΔCt**	**Normalized Copy Number**	**Ø Copy Number**	**StDev**
pK18mob2-*repABC*$_{pSymA}$	1	20.06	0.06	22.96	0.22	2.90	2.90	0	1	-	-
pK18mob2-*repABC*$_{pSymA}$	0.1	23.13	0.36	25.75	0.23	2.62	2.62	0	1		
pLoriV-Mla R1	1	20.45	0.25	23.46	0.01	3.01	2.90	0.11	0.9	0.9	0.17
pLoriV-Mla R1	0.1	24.54	0.84	27.02	0.12	2.48	2.62	-0.15	1.1		
pLoriV-Mla R2	1	21.10	0.04	24.32	0.32	3.23	2.90	0.33	0.8		
pLoriV-Mla R2	0.1	24.26	0.27	27.34	0.12	3.08	2.62	0.46	0.7		
pLoriV-Ti R1	1	21.33	0.06	24.07	0.25	2.74	2.90	-0.16	1.1	1.0	0.21
pLoriV-Ti R1	0.1	24.51	0.11	27.72	0.25	3.21	2.62	0.59	0.7		
pLoriV-Ti R2	1	21.27	0.17	24.04	0.13	2.78	2.90	-0.13	1.1		
pLoriV-Ti R2	0.1	24.53	0.14	27.17	0.39	2.64	2.62	0.02	1.0		
Strain	**[DNA]**	**Ct (Chr)**	**StDev (Ct(Chr))**	**Ct (mob)**	**StDev (Ct(*mob*))**	**ΔCt**	**ΔCt ref**	**ΔΔCt**	**Normalized Copy Number**	**Ø Copy Number**	**StDev**
pK18mob2-*repABC*$_{pSymA}$	1	20.33	0.10	22.33	0.32	2.00	2.00	0	1	-	-
pK18mob2-*repABC*$_{pSymA}$	0.1	23.26	0.43	25.54	0.18	2.28	2.28	0	1		
pLoriV-Mlb R1	1	20.85	0.28	23.21	0.10	2.36	2.00	0.36	0.8	0.8	0.15
pLoriV-Mlb R1	0.1	23.71	0.08	26.00	0.20	2.29	2.28	0.01	1.0		
pLoriV-Mlb R2	1	21.78	0.03	24.44	0.36	2.66	2.00	0.66	0.6		
pLoriV-Mlb R2	0.1	24.72	0.19	27.18	0.02	2.46	2.28	0.18	0.9		
pLoriV-42f2 R1	1	21.14	0.05	24.09	0.25	2.95	2.00	0.95	0.5	0.5	0.06
pLoriV-42f2 R1	0.1	24.29	0.09	27.76	0.18	3.47	2.28	1.19	0.4		
pLoriV-42f2 R2	1	20.97	0.40	24.29	0.41	3.32	2.00	1.32	0.4		
pLoriV-42f2 R2	0.1	24.23	0.13	27.49	0.37	3.25	2.28	0.97	0.5		
pLoriV-42f1 R1	1	20.91	0.19	24.99	0.05	4.08	2.00	2.07	0.2	0.2	0.04
pLoriV-42f1 R1	0.1	24.21	0.17	28.62	0.46	4.41	2.28	2.13	0.2		
pLoriV-42f1 R2	1	20.76	0.06	25.33	0.42	4.56	2.00	2.56	0.2		
pLoriV-42f1 R2	0.1	24.30	0.54	28.49	0.16	4.20	2.28	1.92	0.3		

Cell lysates of *S. meliloti* Sma818 carrying pK18mob2-*repABC*$_{pSymA}$ and *S. meliloti* Rm1021 carrying pLoriV-42b, -42d, -p42e, -p42f1, -p42f2, -pMla, -pMlb and –pTi were obtained from exponentially growing cultures. DNA concentration of cell lysates was adjusted applying Qubit™ fluorometric quantitation. The copy number was determined from the ratio of a chromosomal sequence (Chr; primers 78 and 79) to a target sequence within the *mob* site (*mob*, primers 252 and 253) and *S. meliloti* Rm1021 served as reference strain (grey background). StDev: standard deviation; R1/2: biological replicates. [DNA]: DNA concentration determined by fluorometric quantitation. ref: Reference.

Tabel S7. Detailed description of pABCs generated in this study.

pABC	Construction
pABCa	The oriVSm module part comprising *repABC*$_{pMIb}$ was PCR-amplified from pJD50 using primers 141 and 142. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{p15A}$ was PCR-amplified from pACYC184 using primers 333 and 334. The AR module part comprising P*min2- aacC1* was PCR-amplified from pMW17 using primers 284 and 247. The synTer-MCS module part (containing *synTer*1) was released via digestion of pJD104 using SmaI. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides 340, 341, 344 and 345.
pABCb	The oriVSm module part comprising *repABC*$_{p42b}$ was PCR-amplified from pJD84B using primers 211 and 212. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pMB1}$ was PCR-amplified from pK18mob2 using primers 335 and 336. The AR module part comprising P*min2- aadA1* was PCR-amplified from pMW18 using primers 284 and 249. The synTer-MCS module part (containing *synTer*2) was released via digestion of pJD110 using SmaI. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides 346, 347, 350 and 351.
pABCc	The oriVSm module part comprising *repABC*$_{p42d}$ was PCR-amplified from pJD85A using primers 215 and 216. The oriVEc module part comprising *E.coli* replication origin *oriV*$_{pSC101*}$ was PCR-amplified from pXG-10 using primers 338 and 380. The AR module part comprising P*min2- tmpR* was PCR-amplified from pMW19 using primers 284 and 259. The synTer-MCS module part (containing *synTer*3) was released via digestion of pJD111 using *Sma*I. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides 352, 353, 382 and 383. Antibiotic resistance cassette of *de novo* constructed pABC-replicon was subtituted with a AR module part comprising P*min2-nptII* (PCR-amplified from pMW39 using primers 284 and 261) via AseI and BamHI restriction sites.
pABCc-mob	The oriVSm module part comprising *repABC*$_{p42d}$ was PCR-amplified from pJD85A using primers 215 and 216. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pSC101*}$ was PCR-amplified from pXG-10 using primers 338 and 380. The AR module part comprising P*min2- tmpR* was PCR-amplified from pMW19 using primers 284 and 259. The oriT module part comprising *mob* was PCR-amplified from pK18mob2 using primers 331 and 332. The synTer-MCS module part (containing *synTer*3) was released via digestion of pJD111 using SmaI. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides 352, 353, 354, 381 and 382. The antibiotic resistance cassette of the *de novo* constructed pABC-replicon was subtituted with an AR module part comprising P*min2-nptII* (PCR-amplified from pMW39 using primers 284 and 261) via AseI and BamHI restriction sites.
pABC1	The oriVSm module part comprising *repABC*$_{p42b}$ was PCR-amplified from pLoriV-42b using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pMB1}$ was PCR-amplified from pK18mob2 using primers 546 and 547. The AR module part comprising P*min2-aadA1* was PCR-amplified from pLAR-Spec using primers 502 and 503. The synTer-MCS module part (containing *synTer*1) was PCR-amplified from pLsynTer-1 using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.
pABC1-T2	The oriVSm module part comprising *repABC*$_{p42b}$ was PCR-amplified from pLoriV-42b using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin pMB1 was PCR-amplified from pK18mob2 using primers 546 and 547. The AR module part comprising P*min2-aadA1* was PCR-amplified from pLAR-Spec using primers 502 and 503. The synTer-MCS module part (containing *synTer*2) was PCR-amplified from pLsynTer-2 using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.
pABC1-T3	The oriVSm module part comprising *repABC*$_{p42b}$ was PCR-amplified from pLoriV-42b using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pMB1}$ was PCR-amplified from pK18mob2 using primers 546 and 547. The AR module part comprising P*min2-aadA1* was PCR-amplified from pLAR-Spec using primers 502 and 503. The synTer-MCS module part (containing *synTer*3) was PCR-amplified from pLsynTer-1 using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.
pABC1-loxA	The oriVSm module part comprising *repABC*$_{p42b}$ was PCR-amplified from pLoriV-42b using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pMB1}$ was PCR-amplified from pK18mob2 using primers 546 and 547. The AR module part comprising P*min2-aadA1* was PCR-amplified from pLAR-Spec using primers 502 and 503. The oriT module part comprising *mob-loxRE* was PCR-amplified from pLoriT-1-lox using primers 500 and 501. The synTer-MCS module part (containing *synTer*1-*loxLE*) was PCR-amplified from pLsynTer-1-lox using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3a, BO3b and BO4.
pABC1-loxB	The oriVSm module part comprising *repABC*$_{p42b}$ was PCR-amplified from pLoriV-42b using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pMB1}$ was PCR-amplified from pK18mob2 using primers 546 and 547. The AR module part comprising P*min2-aadA1* was PCR-amplified from pLAR-Spec using primers 502 and 503. The oriT module part comprising *mob-loxRE- rrnBT2* was PCR-amplified from pLoriT-1-loxT using primers 500 and 501. The synTer-MCS module part (containing *synTer*1-*loxLE*) was PCR-amplified from pLsynTer-1-lox using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3a, BO3b and BO4.
pABC2	The oriVSm module part comprising *repABC*$_{pMIa}$ was PCR-amplified from pLoriV-MIa using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{p15A}$ was PCR-amplified from pACYC184 using primers 544 and 545. The AR module part comprising P*min2-aacC1* was PCR-amplified from pLAR-Gm using primers 502 and 503. The synTer-MCS module part (containing *synTer*2) was PCR-amplified from pLsynTer-2 using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.
pABC3	The oriVSm module part comprising *repABC*$_{pTi}$ was PCR-amplified from pLoriV-Ti using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pSC101*}$ was PCR-amplified from pXG-10 using primers 670 and 671. The AR module part comprising P*min2-hph* was PCR-amplified from pLAR-Hyg using primers 502 and 503. The synTer-MCS module part (containing *synTer*3) was PCR-amplified from pLsynTer-3 using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.
pABC4	The oriVSm module part comprising *repABC*$_{p42e}$ was PCR-amplified from pLoriV-42e using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pMB1}$ was PCR-amplified from pK18mob2 using primers 546 and 547. The AR module part comprising P*min2-aadA1* was PCR-amplified from pLAR-Spec using primers 502 and 503. The synTer-MCS module part (containing *synTer*1) was PCR-amplified from pLsynTer-1 using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.
pABC5	The oriVSm module part comprising *repABC*$_{p42f1}$ was PCR-amplified from pLoriV-42f1 using primers 494 and 495. The oriVEc module part comprising *E. coli* replication origin *oriV*$_{pMB1}$ was PCR-amplified from pK18mob2 using primers 546 and 547. The AR module part comprising P*min2-aadA1* was PCR-amplified from pLAR-Spec using primers 502 and 503. The synTer-MCS module part (containing *synTer*2) was PCR-amplified from pLsynTer-2 using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.
pABC6	The oriVSm module part comprising *repABC*$_{p42f2}$ was PCR-amplified from pLoriV-42f2 using primers 494 and 495. The oriVEc module part comprising *E.coli* replication origin *oriV*$_{pMB1}$ was PCR-amplified from pK18mob2 using primers 546 and 547. The AR module part comprising P*min2-aadA1* was PCR-amplified from pLAR-Spec using primers 502 and 503. The synTer-MCS module part (containing *synTer*2) was PCR-amplified from pLsynTer-2 using primers 496 and 497. Module parts were assembled via ligase cycling reaction using bridging oligonucleotides BO1, BO2, BO3 and BO4.

Table S8. Flow cytometry measurements of *S. meliloti* Rm1021 carrying pABC-eGFP derivatives.

eGFP	0 h				24 h				48 h				96 h				96 h stationary			
Run 1	R1	R2	R3	**mean**	R1	R2	R3	**mean**	R1	R2	R3	**mean**	R1	R2	R3	**mean**	R1	R2	R3	**mean**
***S. meliloti* MWSm37**	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**
***S. meliloti* Rm1021**	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**
S. meliloti*/pABCa-*eGFP	100	100	100	**100**	99	99	99	**99**	99	99	99	**99**	97	97	97	**97**	98	98	98	**98**
S. meliloti*/pABCb-*eGFP	100	100	100	**100**	99	99	100	**100**	99	99	99	**99**	99	98	99	**99**	99	99	99	**99**
S. meliloti*/pABCc-mob-*eGFP	100	100	100	**100**	99	99	99	**99**	99	98	99	**99**	98	97	97	**97**	98	98	98	**98**
3xpABC$^{a\text{-}eGFP}$	100	100	100	**100**	99	99	99	**99**	98	98	98	**98**	95	97	96	**96**	97	97	98	**97**
3xpABC$^{b\text{-}eGFP}$	100	100	100	**100**	99	99	99	**99**	99	99	99	**99**	98	98	98	**98**	98	98	98	**98**
3xpABC$^{c\text{-}eGFP}$	100	100	100	**100**	99	99	99	**99**	99	99	99	**99**	98	97	97	**97**	98	98	98	**98**
Run 2																				
S. meliloti MWSm37	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**
S. meliloti Rm1021	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**
pABC1-eGFP	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**	99	99	99	**99**	100	100	100	**100**
pABC2-eGFP	100	100	100	**100**	81	85	80	**82**	51	56	49	**52**	13	13	13	**13**	35	37	37	**36**
pABC3-eGFP	100	100	94	**-**	99	99	36	**-**	49	99	2	**-**	1	93	0	**-**	4	98	0	**-**
pABC4-eGFP	100	100	100	**100**	99	99	99	**99**	98	98	98	**98**	97	98	97	**97**	98	98	98	**98**
2xpABC$^{3\text{-}eGFP}$	99	99	100	**-**	88	99	99	**-**	7	99	98	**-**	2	97	78	**-**	3	98	96	**-**
Run 3																				
Positive (MWSm37)	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**
S. meliloti Rm1021	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**
2xpABC$^{4\text{-}eGFP}$	99	100	99	**99**	99	99	99	**99**	98	98	98	**98**	97	97	97	**97**	98	98	98	**98**
S. meliloti*/pPHU231-*eGFP	98	97	98	**98**	80	80	80	**80**	63	63	63	**63**	39	39	39	**39**	54	55	55	**55**

***eGFP*-expR**	0 h				24 h				48 h				96 h			
Run 1	R1	R2	R3	**mean**	R1	R2	R3	**mean**	R1	R2	R3	**mean**	R1	R2	R3	**mean**
S. meliloti MWSm37	99	100	100	**99**	100	100	100	**100**	100	100	100	**100**	100	100	100	**100**
S. meliloti Rm1021	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**	0	0	0	**0**
S. meliloti/pABCa-eGFP-expR	100	100	100	**100**	98	98	98	**98**	95	95	95	**95**	84	84	82	**83**
S. meliloti/pABCb-eGFP-expR	100	100	100	**100**	98	98	98	**98**	97	98	97	**97**	94	94	94	**94**
S. meliloti/pABCc-eGFP-expR	99	100	100	**100**	99	99	99	**99**	98	98	98	**98**	93	91	91	**92**
S. meliloti/pABC2-eGFP-expR	100	100	100	**100**	67	66	66	**66**	17	17	17	**17**	1	1	1	**1**
S. meliloti/pABC4-eGFP-expR	99	100	99	**99**	98	97	98	**98**	92	92	92	**92**	66	65	58	**63**

Three replicates were measured (R1-3) after 0, 24 and 96 h post-inoculation. Strains MWSm37 and Rm1021 served as positive and negative controls, respectively, and verified valid measurements for every run. 3xpABC$^{a\text{-}eGFP}$: *S. meliloti* carrying pABCa-eGFP, pABCb and pABC-mob; 3xpABC$^{b\text{-}eGFP}$: *S. meliloti* carrying pABCa, pABCb-eGFP and pABC-mob; 3xpABC$^{c\text{-}eGFP}$: *S. meliloti* carrying pABCa, pABCb and pABC-mob-eGFP. 2xpABC$^{3\text{-}eGFP}$: *S. meliloti* carrying pABC3-eGFP and pABC4. 2xpABC$^{4\text{-}eGFP}$: *S. meliloti* carrying pABC3 and pABC4-eGFP.

Table S9 (1/2). Strains and plasmids used in the supplementary information.

Strains/ Plasmids	Description	References
S. meliloti		
Sm-2738	Transposon mutant strain carrying a trimethoprim resistance cassette (*tmpR*) derived from plasmid R751 from *Enterobacter aerogenes* (Tmp^r)	A. Becker
BO33	Rm1021 carrying excision module for Cre/*loxP* studies (Gm^r)	Harrison et al., 2011
JDSm15	Strain Rm1021 bearing pJD10 integrated into pSymB (Km^r)	This study
JDSm16	Strain Rm1021 bearing pJD13 integrated into pSymB (Km^r)	This study
JDSm17	Strain Rm1021 bearing pJD14 integrated into pSymB (Km^r)	This study
MWSm40	Strain ΔhsdR bearing pJD112 integrated into pSymB (Km^r)	This study
MWSm41	Strain ΔhsdR bearing pJD113 integrated into pSymB (Km^r)	This study
MWSm42	Strain ΔhsdR bearing pJD114 integrated into pSymB (Km^r)	This study
MWSm9	Strain ΔhsdR bearing pMW1 integrated into pSymB (Km^r)	This study
MWSm43	Strain ΔhsdR bearing pMW40 integrated into pSymB (Km^r)	This study
MWSm36	Strain ΔhsdR bearing pMW27 integrated into pSymB (Km^r)	This study
MWSm38	Strain ΔhsdR bearing pMW29 integrated into pSymB (Km^r)	This study
MWSm39	Strain ΔhsdR bearing pJD10 integrated into pSymB (Km^r)	This study
MWSm26	Strain ΔhsdR bearing pMW17 integrated into pSymB (Km^r, Gm^r)	This study
MWSm27	Strain ΔhsdR bearing pMW18 integrated into pSymB (Km^r, $Spec^r$)	This study
MWSm65	Strain ΔhsdR bearing pMW39 integrated into pSymB (Km^r, Gm^r)	This study
MWSm63	Strain ΔhsdR bearing pMW37 integrated into pSymB (Km^r, Tc^r)	This study
MWSm64	Strain ΔhsdR bearing pMW38 integrated into pSymB (Km^r, Hyg^r)	This study
MWSm28	Strain ΔhsdR bearing pMW19 integrated into pSymB (Km^r, Tmp^r)	This study
JDSm150	Strain SmCreΔhsdR carrying pABC1-loxA ($Spec^r$)	This study
JDSm151	Strain SmCreΔhsdR carrying pABC1-loxB ($Spec^r$)	This study
JDSm159	Strain SmCreΔhsdR carrying pABC1-loxA after Cre/*lox* mediated deletion of modules oriVEc and oriT ($Spec^r$)	This study
JDSm160	Strain SmCreΔhsdR carrying pABC1-loxB after Cre/*lox* mediated deletion of modules oriVEc and oriT ($Spec^r$)	This study
JDSm201	Strain SmCreΔhsdR bearing P*min2-loxR-aacC1* downstream of the *exo* gene cluster; precursor of SmCre_exo-IN (Gm^r)	This study
Plasmids		
pMS252	pUC derivative carrying a gentamicin resistance cassette derived from RIO33 [TnI696]	Becker et al.,1995
pLK64	pPHU231-*eGFP* containing *sinI* promoter	McIntosh et al.,2008
pLAU53	pBAD24 derivate carrying *tetR-YFP* translational fusion	Lau et al., 2003
pAR46	Providing the *parS* site derived from *Yersinia pestis* ($parS_{Yp}$)	A. Raßbach & M. Thanbichler, unpublished
pAR51	Providing a N-terminal fusion of *mChr* to *Yersinia pestis* derived *parB* ($mChr\text{-}parB_{Yp}$)	A. Raßbach & M. Thanbichler, unpublished
pSRKKm/*expR*	pSRKKm carrying *expR* under control of P*lacZ*	M. McIntosh & A. Becker, unpublished
pCBT113	Suicide vector carrying hygromycin resistance gene *hph*	Sauviac et al., 2007
pJD10	pK18mob2 derivative carrying P*aacC1-eGFP*	This work
pJD13	pK18mob2 derivative carrying P*aacC1-rrnBT1-eGFP*	This work
pJD14	pK18mob2 derivative carrying P*aacC1-rrnBT2-eGFP*	This work

Asterisks indicate intermediate constructs only used for constructions. Km^r/$Spec^r$/Tc^r/Hyg^r/Gm^r/Tmp^r: Kanamycin/spectinomycin/tetracycline/ hygromycin/gentamicin/trimethoprim resistance.

Harrison, C. L., Crook, M. B., Peco, G., Long, S. R., and Griffitts, J. S. (2011) Employing site-specific recombination for conditional genetic analysis in *Sinorhizobium meliloti*. *Appl. Environ. Microbiol. 77, 3916–3922*.

Becker, A., Schmidt, M., Jäger, W., and Pühler, A. (1995) New gentamicin-resistance and *lacZ* promoter-probe cassettes suitable for insertion mutagenesis and generation of transcriptional fusions. *Gene 162, 37–39*.

McIntosh, M., Krol, E., and Becker, A. (2008) Competitive and Cooperative Effects in Quorum-Sensing-Regulated Galactoglucan Biosynthesis in *Sinorhizobium meliloti*. *J. Bacteriol.* 190, 5308–5317.

Lau, I. F., Filipe, S. R., Søballe, B., Økstad, O.-A., Barre, F.-X., and Sherratt, D. J. (2003) Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol. 49, 731–743*.

Sauviac, L., Philippe, H., Phok, K., and Bruand, C. (2007) An extracytoplasmic function sigma factor acts as a general stress response regulator in *Sinorhizobium meliloti*. *J Bacteriol. 189*, 4204–4216.

Table S9 (2/2). Strains and plasmids used in the supplementary information.

Plasmids		
pJD47*	pK18mob2sacB derivative carrying *thrA-rpoC – loxP – aadA1*	This work
pJD104*	pK18mob2 derivative carrying *synTer1*	This work
pJD110*	pK18mob2 derivative carrying *synTer2*	This work
pJD111*	pK18mob2 derivative carrying *synTer3*	This work
pJD112	Integrative pK18mob2 derivative carrying P*aacC1-synTer1-eGFP*	This work
pJD113	Integrative pK18mob2 derivative carrying P*aacC1-synTer2-eGFP*	This work
pJD114	Integrative pK18mob2 derivative carrying P*aacC1-synTer3-eGFP*	This work
pJD137*	pK18mob2 derivative carrying a synthetic DNA fragment comprising all module linker sequences	This work
pMW3*	pK18mob2 derivative carrying P*min1*	This work
pMW14*	pK18mob2 derivative carrying P*min2*	This work
pMW15*	pK18mob2 derivative carrying P*min3*	This work
pMW40	pK18mob2 derivative carrying promoterless *SD-eGFP* fragment	This work
pMW27	pK18mob2 derivative carrying P*min1-eGFP*	This work
pMW29	pK18mob2 derivative carrying P*min3-eGFP*	This work
pMW17	pK18mob2 derivative carrying P*min2-aacC1*	This work
pMW18	pK18mob2 derivative carrying P*min2-aadA1*	This work
pMW39	pK18mob2 derivative carrying P*min2-nptII*	This work
pMW37	pK18mob2 derivative carrying P*min2-tetA*	This work
pMW38	pK18mob2 derivative carrying P*min2-hph*	This work
pMW19	pK18mob2 derivative carrying P*min2-tmpR*	This work
pABCa-eGFPf	pABCa carrying a forwardly integrated *SD-eGFP* fragment	This work
pABCa-eGFPr	pABCa carrying a reversely integrated *SD-eGFP* fragment	This work
pABCb-eGFPf	pABCb carrying a forwardly integrated *SD-eGFP* fragment	This work
pABCb-eGFPr	pABCb carrying a reversely integrated *SD-eGFP* fragment	This work
pABCb-tetO	pABCb carrying a 4.1 kb *tetO* array	This work
pABCb-eGFP	pABCb carrying a constitutive *eGFP* cassette	This work
pABCc-mob-eGFPf	pABCc-mob carrying a forwardly integrated *SD-eGFP* fragment	This work
pABCc-mob-eGFPr	pABCc-mob carrying a reversely integrated *SD-eGFP* fragment	This work
pABC1-tetO	pABC1 carrying a 4.1 kb *tetO* array	This work
pABC2-tetO	pABC2 carrying a 4.1 kb *tetO* array	This work
pABC4-tetO	pABC1 carrying a 4.1 kb *tetO* array	This work
pABC5-parS_{Yp}	pABC5 carrying a *Yersinia pestis* derived *parS* sequence (*parS_{Yp}*)	This work
pABC6-parS_{Yp}	pABC6 carrying a *Yersinia pestis* derived *parS* sequence (*parS_{Yp}*)	This work

Asterisks indicate intermediate constructs only used for constructions.

Tabel S10 (1/3). Detailed description of constructs generated in this study.

Construct	Construction
pMla_s	pMla derived *repABC* cassette was PCR amplified with primers 139+140 from *M. loti* MAFF303099 cell lysate and reversely integrated into SmaI-digested pK18mob2.
p42e_s	p42e derived *repABC* cassette was PCR amplified with primers 145+146 from *R. etli* CFN42 cell lysate and reversely integrated into SmaI-digested pK18mob2.
pJD10	A homologous region (HR) for integration at Smb1168127 was PCR amplified with primers 21+22 from genomic DNA of *S. meliloti* Rm1021 and inserted into SmaI-opened pK18mob2. P*aacC1* was PCR amplified with primers 6+7 from pMS252. A *loxP* site (which is not relevant to this construct and just acts as spacer sequence) was generated via hybridization of primers 15+16. A promoterless *eGFP* fragment (containing a SD sequence) was amplified with primers 2 and 3 from pLK64. Finally, BamHI/NdeI-cut P*aacC1*, NdeI/PstI-cut *loxP* and PstI/HindIII-digested *eGFP* was ligated into BamHI/HindIII-opened pK18mob2-HR.
pJD13	*rrnBT1* was generated via hybridization of primers 57-60. The resulting fragment was digested with SalI and PstI and inserted into SalI/PstI-opened pJD10.
pJD14	*rrnBT2* was generated via hybridization of primers 61-64. The resulting fragment was digested with NdeI and SalI and inserted into NdeISalI-opened pJD10.
pJD47	*rpoC-thrA* was PCR amplified with primers 131+132 from *S. meliloti* BO33. Generation of *loxP* via hybridization of primers 15 and 16. PCR amplification of *aadA1* with primers 127+128 (carrying an additional ScaI recognition site). Ligation of BamHI/NdeI-digested *rpoC-thrA*, NdeI/SalI-digested *loxP* and SalI/PstI-digested *aadA1* into BamHI/PstI-opened pK18mobsacB.
pJD104	Synthetic DNA fragment *synTer1_pre* was digested with SalI and integrated into SmaI/SalI-opened pK18mob2. *rrnBT2* was generated via hybridization of primers 301/302 and inserted into SalI/SphI-opened pK18mob2-*synTer1_pre*.
pJD110	Synthetic DNA fragment *synTer2_pre* was digested with SalI and integrated into SmaI/SalI-opened pK18mob2. *rrnBT2* was generated via hybridization of primers 301/302 and inserted into SalI/SphI-opened pK18mob2-*synTer2_pre*.
pJD111	Synthetic DNA fragment *synTer3_pre* was digested with SalI and integrated into SmaI/SalI-opened pK18mob2. *rrnBT2* was generated via hybridization of primers 301/302 and inserted into SalI/SphI-opened pK18mob2-*synTer3_pre*.
pJD112	*synTer1* was released from NdeI/PstI-digested pJD104 and ligated into NdeI/PstI-opened pJD10.
pJD113	*synTer2* was released from NdeI/PstI-digested pJD110 and ligated into NdeI/PstI-opened pJD10.
pJD114	*synTer3* was released from NdeI/PstI-digested pJD111 and ligated into NdeI/PstI-opened pJD10.
pJD120	A homologous region comprising the upstream region of *tauA* (HR1, Smb1050453-1050952) was PCR amplified with primers 266 and 267 from *S. meliloti* Rm1021 genomic DNA. A homologous region comprising P*tauAB-tauA* (HR2, Smb1050781-1051458) was PCR amplified with primers 269 and 330. EcoRI/SalI-digested HR1 and SalI/SphI-digested HR2 were ligated into EcoRI/SphI-opened pK18mobsacB. *tetR-YFP* was PCR-amplified with primers 270+271 from pLAU53. *mCHr-parB*$_{Yp}$ was PCR-amplified with primers 272 and 273 from pAR51. SalI/NheI-digested *tetR-YFP* and NheI/SphI-digested *mCHr-parB*$_{Yp}$ were inserted into SalI/SphI-opened pK18mobsacB. The resulting construct was digested with SalI/ScaI thereby releasing *tetR-YFP-mCHr-parB*$_{Yp}$ which was ligated into SalI/ScaI-opened pK18mobsacB-HR1-HR2 (in between the homologous regions).
pJD137	Synthetic DNA fragment *module_linkers* comprising all module linkers was digested with HindIII and inserted into SmaI/HindIII-opened pK18mob2.
pJD211	P*min2* (lacking a SD sequence) was generated via hybridization of primers 388/389. *loxR* was generated via hybridization of primers 189/190. *aacC1* was PCR amplified with primers 125 and 126 from pMS252. Hybridization products P*min2* and *loxR*, and SalI/PstI-digested *aacC1* were integrated into BamHI/PstI-opened pK18mob2.
pMW1	The homologous region for integration at Smb1166733 (*exsA*) was PCR-amplified from *S. meliloti* Rm1021 genomic DNA using primers 94 + 95 and cut with EcoRI/BamHI. The fragment was integrated into EcoRI/BamHI-cut pK18mob2.
pMW3	Promoter region of P*min1* was generated via hybridization of primers 244 and 245. *loxLR* site was generated via hybridization of primers 191 and 192. The fragments were triple-ligated into SphI/SalI-cut pK18mob2. The homologous region for integration at Smb1166733 (*exsA*) was PCR-amplified from *S. meliloti* Rm1021 genomic DNA using primers 250 + 251 and cut with KpnI/EcoRI. The fragment was integrated into KpnI/EcoRI-cut pK18mob2 comprising P*min1* and *loxLERE*.
pMW14	In order to substitute promoter region P*min1* with P*min2*, pMW3 was digested with enzymes SphI and NdeI. Subsequently, 5' phosphorylated and hybridized primers 284 and 285 were integrated.
pMW15	In order to substitute promoter region P*min1* with P*min3* 5' phosphorylated and hybridized oligonucleotides 286 and 287 as well as 288 and 289 were triple-ligated into SphI/NdeI-cut pMW3.
pMW40	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW28 using primers 366 + 327 and digested with SphI/BamHI. The fragment was integrated into SphI/BamHI-cut pMW14.
pMW27	*eGFP* was PCR-amplified from pJD10 using primers 326 + 327 and digested with NdeI/BamHI. The fragment was integrated into NdeI/BamHI-cut pMW3.
pMW28	*eGPF* was PCR-amplified from pJD10 using primers 326 + 327 and digested with NdeI/BamHI. The fragment was integrated into NdeI/BamHI-cut pMW14.
pMW29	*eGFP* was PCR-amplified from pJD10 using primers 326 + 327 and digested with NdeI/BamHI. The fragment was integrated into NdeI/BamHI-cut pMW15.
pMW17	*aacC1* was amplified from *S. meliloti* B033 genomic DNA using primers 290 + 247 and digested with NdeI/BamHI. The fragment was integrated into NdeI/BamHI-cut pMW14.
pMW18	*aadA1* was amplified from *S. meliloti* B033 genomic DNA using primers 291 + 249 and digested with NdeI/BamHI. The fragment was integrated into NdeI/BamHI-cut pMW14.
pMW39	Construct pMW3 was digested with enzymes HindIII/EcoRI and the released fragment was integrated into HindIII/EcoRI-cut pG18mob2. The resulting construct was digested with enzymes SphI/BamHI. Subsequently, *nptII* was PCR-amplified from pK18mob2 with primers 369 + 261 and digested with NdeI/BamHI. The PCR product and hybridized primers 284 and 285 (P*min2*) were triple-ligated via SphI/BamHI.
pMW37	*tetA* was amplified from pACYC184 using primers 367 + 436 and digested with NdeI/KpnI. The fragment was integrated into NdeI/KpnI-cut pMW14.
pMW38	*hph* was amplified from pCBT113 using primers 368 + 255 and digested with NdeI/BamHI. The fragment was integrated into NdeI/BamHI-cut pMW14.
pMW19	*tmpR* was amplified from *S. meliloti* strain Sm-2738 using primers 292 + 259 and digested with NdeI/BamHI. The fragment was integrated into NdeI/BamHI-cut pMW14.
pPHU231-eGFP	An *eGPF* cassette including P*min2* was PCR-amplified from pMW28 using primers 647 + 648 and digested with KpnI/EcoRI. The fragment was inserted into KpnI/EcoRI-cut pPHU231.
pABCa-eGFP	An *eGPF* cassette including P*min2* was PCR-amplified from pMW28 using primers 284 + 327. The fragment was integrated in forward orientation into ScaI-cut pABCa.
pABCa-eGFP-expR	An *eGPF* cassette including P*min2* and *expR* including P*lacZ* promoter was PCR-amplified from pABC1-eGFP-expR using primers 651 + 652 and digested with PacI/ScaI. The fragment was integrated into PacI/ScaI-cut pABCa.
pABCa-eGFPf	Sequence of promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in forward orientation into ScaI-cut pABCa.

Tabel S10 (2/3). Detailed description of constructs generated in this study.

Construct	Construction
pABCa-eGFPr	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in reverse orientation into ScaI-cut pABCa.
pABCa-tetO	4.2 kb *tetO* array was released by digestion of pLAU44 with EcoRV and HincII. The fragment was integrated into ScaI-cut pABCa.
pABCb-eGFP	*eGPF* including promoter P*min2* was PCR-amplified from pMW28 using primers 284 + 327. The fragment was integrated in forward orientation into ScaI-cut pABCb.
pABCb-eGFP-expR	*eGFP* including P*min2* and *expR* including a P*lacZ* promoter was PCR-amplified from pABC1-*eGFP-expR* using primers 651 + 652 and digested with PacI/ScaI. The fragment was integrated into PacI/ScaI-cut pABCb.
pABCb-eGFPf	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in forward orientation into ScaI-cut pABCb.
pABCb-eGFPr	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in reverse orientation into ScaI-cut pABCb.
pABCb-tetO	4.2 kb *tetO*-array was released via digestion of pLAU44 with EcoRV and HincII. The fragment was integrated into ScaI-cut pABCb.
pABCb-parS_{Yp}	The *parS* site from *Yersinia pestis* was amplified from pAR46 using primers 448 + 449 and digested with NotI. The fragment was integrated into NotI-cut pABCb.
pABCc-eGFP-expR	*eGFP* including P*min2* promoter and *expR* including P*lacZ* promoter was PCR-amplified from pABC1-eGFP-expR using primers 651 + 652 and digested with PacI/ScaI. The fragment was integrated into PacI/ScaI-cut pABCc.
pABCc-parS_{Yp}	*parS* site from *Yersinia pestis* was amplified from pAR46 using primers 448 + 449 and digested with NotI. The fragment was integrated into NotI-cut pABCc.
pABCc-mob-eGFP	*eGPF* including promoter P*min2* was PCR-amplified from pMW28 using primers 284 + 327. The fragment was integrated in forward orientation into ScaI-cut pABCc-mob.
pABCc-mob-eGFPf	Promoterless *eGFP* including SD of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in forward orientation into ScaI-cut pABCc-mob.
pABCc-mob-eGFPr	Promoterless *eGFP* including SD of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in reverse orientation into ScaI-cut pABCc-mob.
pABC1-eGFP	*eGPF* including promoter P*min2* was PCR-amplified from pMW28 using primers 651 + 648 and digested with PacI/NotI. The fragment was integrated into PacI/NotI-cut pABC1.
pABC1-eGFP-expR	*eGPF* including promoter P*min2* was PCR-amplified from pMW28 using primers 651 + 648 and digested with PacI/NotI. Sequence of *expR* including P*lacZ* promoter and native SD was PCR-amplified from pSRKKm-expR using primers 649 + 652 and digested with NotI/ScaI. The fragments were triple-ligated into PacI/ScaI-cut pABC1.
pABC1-tetO	4.2 kb *tetO* array was released via digestion of pLAU44 with EcoRV and HincII. The fragment was integrated into PmeI-cut pABC1.
pABC1-eGFPf	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in forward orientation into ScaI-cut pABC1.
pABC1-eGFPr	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in reverse orientation into ScaI-cut pABC1.
pABC1-T2-eGFPf	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in forward orientation into ScaI-cut pABC1-T2.
pABC1-T2-eGFPr	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in reverse orientation into ScaI-cut pABC1-T2.
pABC1-T3-eGFPf	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in forward orientation into ScaI-cut pABC1-T3.
pABC1-T3-eGFPr	Promoterless *eGFP* including the SD sequence of P*min2* was PCR-amplified from pMW40 using primers 366 + 327. The fragment was integrated in reverse orientation into ScaI-cut pABC1-T3.
pABC2-eGFP	*eGPF* including P*min2* was PCR-amplified from pMW28 using primers 651 + 648 and digested with PacI. The fragment was integrated into PacI/ScaI-cut pABC2.
pABC2-eGFP-expR	*eGFP* including P*min2* and *expR* including P*lacZ* was PCR-amplified from pABC1-*eGFP-expR* using primers 651 + 652 and digested with PacI/ScaI. The fragment was integrated into PacI/ScaI-cut pABC2.
pABC2-tetO	4.2 kb *tetO* array was released via digestion of pLAU44 with EcoRV and HincII. The fragment was integrated into PmeI-cut pABC2.
pABC3-eGFP	*eGPF* including P*min2* was PCR-amplified from pMW28 using primers 651 + 648 and digested with PacI/NotI. The fragment was integrated into PacI/NotI-cut pABC3.
pABC4-eGFP	*eGPF* including P*min2* was PCR-amplified from pMW28 using primers 651 + 648 and digested with PacI/NotI. The fragment was integrated into PacI/NotI-cut pABC4.
pABC4-eGFP-expR	*eGFP* including P*min2* and *expR* including P*lacZ* were PCR-amplified from pABC1-*eGFP-expR* using primers 651 + 652 and digested with PacI/ScaI. The fragment was integrated into PacI/ScaI-cut pABC4.
pABC4-tetO	4.2 kb *tetO* array was released via digestion of pLAU44 with EcoRV and HincII. The fragment was integrated into PmeI-cut pABC4.
pABC5-parS_{Yp}	The *parS* site derived from *Yersinia pestis* was amplified from pAR46 using primers 448 + 449. The fragment was integrated into PmeI-cut pABC5.
pABC6-parS_{Yp}	The *parS* site derived from *Yersinia pestis* was amplified from pAR46 using primers 448 + 449 and digested with NotI. The fragment was integrated into NotI-cut pABC6.
pK18mob2-repABC$_{pSymA}$	The pSymA-derived *repA2B2C2* cassette was PCR amplified with primers 48+50 from *S. meliloti* Rm1021 genomic DNA. The KpnI/BamHI digested PCR product was inserted into KpnI/BamHI-opened pK18mob2.
pABCb-lacI	The P*lacI* promoter sequence comprising the SD sequence from P*min2* was PCR-amplified from pSRKGm using primers 608 + 614 and digested with PacI/NdeI. *lacI* including the ATG start codon was PCR-amplified from pSRKGm using primers 613 + 609 and digested with NdeI/PmeI. The fragments were triple-ligated into PacI/PmeI-cut pABCb.
pMlb_lacO	A *lac* operator (*lacO*) sequence was integrated into the predicted transcription start site of *repABC*$_{pMlb}$ via overlap-PCR using *M. loti* MAFF303099 cell lysate as template and primers 141+629 and 142+630, respectively. Both PCR products served as template for a further PCR using primers 141+142. The *repABC*$_{pMlb}$-*lacO* sequence was finally cloned blunt end into SmaI-opened pK18mob2.
pIso	*loxL* was generated via hybridization of primers 187 and 188 and inserted into NdeI/SalI-opened pJD47 (giving rise to a substitution of *loxP* with *loxL*). A homologous region (HR) for integration at Smb1166733 was PCR amplified with primers 94 and 95 from *S. meliloti* Rm1021 genomic DNA. A fragment spanning *thrA-rpoC–loxL–aadA1* was released from BamHI/PstI-digested pJD47. Then, EcoRI/BamHI-digested HR and the *thrA-rpoC–loxL–aadA1* were inserted into EcoRI/PstI-opened pK18mob2. *repABC*$_{pMlb}$*_lacO* was PCR amplified with primers 141 and 142 from pMlb_lacO and reversely inserted into ScaI-digested pK18mob2-*HR-thrA-rpoC–loxL–aadA1*. The resulting construct was digested with XbaI and ligated with a XbaI-cut *lacI* cassette which was amplified with primers 717 and 718 from pABCb-*lacI*.
pLoriVSm	The linker fragment was generated by hybridization of primers 462/463 and inserted into SmaI/HindIII-opened pK18mob2.
pLsynTer-MCS	A 200 bp fragment was released from SmaI/PstI digested pJD137 and ligated into SmaI/PstI-opened pK18mob2. The resulting construct was digested with XbaI and HindIII. Subsequently, the released backbone was blunted and re-ligated.
pLoriT	A 200 bp fragment was released from SmaI/PstI digested pJD137 and ligated into SmaI/PstI-opened pK18mob2. The resulting construct was digested with SmaI/XbaI. Subsequently, the released backbone was blunted and re-ligated.
pLAR	The backbone of SmaI/PstI digested pJD137 was blunted and re-ligated.
pLoriV-42b	The p42b-derived *repABC* cassette was PCR amplified with primers 211+212 from *R. etli* CFN42 cell lysate and reversely integrated into ScaI/-opened pLoriVSm.
pLoriV-42d	The p42d-derived *repABC* cassette was PCR amplified with primers 215+216 from *R. etli* CFN42 cell lysate and reversely integrated into ScaI/-opened pLoriVSm.
pLoriV-42e	The p42e-derived *repABC* cassette was PCR amplified with primers 145+470 from *R. etli* CFN42 cell lysate and reversely integrated into ScaI/-opened pLoriVSm.
pLoriV-42f1	The p24f-derived *repA1B1C1* cassette was PCR amplified with primers 217+218 from *R. etli* CFN42 cell lysate and reversely integrated into ScaI/-opened pLoriVSm.
pLoriV-42f2	The p42f-derived *repA2B2C2* cassette was PCR amplified with primers 219+220 from *R. etli* CFN42 cell lysate and reversely integrated into ScaI/-opened pLoriVSm.

Tabel S10 (3/3). Detailed description of constructs generated in this study.

Construct	Construction
pLoriV-Mla	The pMla-derived *repABC* cassette was PCR amplified with primers 139+471 from *M. loti* MAFF303099 cell lysate and reversely integrated into ScaI/-opened pLoriVSm.
pLoriV-Mlb	The pMlb-derived *repABC* cassette was PCR amplified with primers 141+142 from *M. loti* MAFF303099 cell lysate and reversely integrated into ScaI/-opened pLoriVSm.
pLoriV-Ti	The pTi-derived *repABC* cassette was PCR amplified with primers 137+138 from *A. tumefaciens* C58 cell lysate and reversely integrated into ScaI/-opened pLoriVSm.
pLsynTer-1	The MCS-containing *synTer1-rrnBT2* was released from SmaI-digested pJD104 and forwardly ligated into ScaI-opened pLoriT.
pLsynTer-2	The MCS-containing *synTer2-rrnBT2* was released from SmaI-digested pJD110 and forwardly ligated into ScaI-opened pLoriT.
pLsynTer-3	The MCS-containing *synTer3-rrnBT2* was released from SmaI-digested pJD111 and forwardly ligated into ScaI-opened pLoriT.
pLsynTer-1-lox	*loxLE* was generated via hybridization of primers 574/575 and inserted into ScaI/NdeI-opened pLsynTer-1.
pLsynTer-2-lox	*loxLE* was generated via hybridization of primers 574/575 and inserted into ScaI/NdeI-opened pLsynTer-2.
pLsynTer-3-lox	*loxLE* was generated via hybridization of primers 574/575 and inserted into ScaI/NdeI-opened pLsynTer-3.
pLoriT-1	A 464 bp fragment comprising a *mob* site was PCR amplified with primers 331 and 332 from pPHU231 and ligated into ScaI-opened pLoriT.
pLoriT-1-lox	*sacB-mob* was PCR amplified with primers 573+576 from pK18mobsacB. *loxRE* was generated via hybridization of primers 581/582 and inserted into ScaI-opened pLoriT. The resulting construct was digested with ScaI/NdeI and ligated with *sacB-mob* (NdeI).
pLoriT-1-loxT	*sacB-mob* was PCR amplified with primers 573+576 from pK18mobsacB. *loxRE-rrnBT2* was generated via hybridization of primers 577/578 and 579/580 and inserted into ScaI-opened pLoriT. The resulting construct was digested with ScaI/NdeI and ligated with *sacB-mob* (NdeI).
pLAR-Km	P*min2-nptII* was PCR amplified with primers 284+261 from pMW39 and forwardly inserted into ScaI-opened pLAR.
pLAR-Spec	P*min2-aadA1* was PCR amplified with primers 284+249 from pMW18 and forwardly inserted into ScaI-opened pLAR.
pLAR-Tc	P*min2-tetR* was PCR amplified with primers 284+436 from pMW37 and forwardly inserted into ScaI-opened pLAR.
pLAR-Hyg	P*min2-hph* was PCR amplified with primers 284+255 from pMW38 and forwardly inserted into ScaI-opened pLAR.
pLAR-Tmp	P*min2-tmpR* was PCR amplified with primers 284+259 from pMW19 and forwardly inserted into ScaI-opened pLAR.
pLAR-Gm	P*min2-aacC1* was PCR amplified with primers 284+247 from pMW17 and forwardly inserted into ScaI-opened pLAR.
p42b_p	*repABC*-based $oriV_{p42b}$ was PCR amplified with primers 211+212 from *R. etli* CFN42 cell lysate. The PCR product was 5‘ phosphorylated and reversely inserted into SmaI-opened pK18mob2.
p42d_p	*repABC*-based $oriV_{p42d}$ was PCR amplified with primers 215+216 from *R. etli* CFN42 cell lysate. The PCR product was 5‘ phosphorylated and reversely inserted into SmaI-opened pK18mob2.
p42f1_p	*repA1B1C1*-based $oriV_{p42f}$ was PCR amplified with primers 217+218 from *R. etli* CFN42 cell lysate. The PCR product was 5‘ phosphorylated and reversely inserted into SmaI-opened pK18mob2.
p42f2_p	*repA2B2C2*-based $oriV_{p42f}$ was PCR amplified with primers 219+220 from *R. etli* CFN42 cell lysate. The PCR product was 5‘ phosphorylated and reversely inserted into SmaI-opened pK18mob2.
pMlb_p	*repABC*-based $oriV_{pMlb}$ was PCR amplified with primers 141+142 from *M. loti* MAFF303099 cell lysate. The PCR product was 5‘ phosphorylated and reversely inserted into SmaI-opened pK18mob2.
pSMED01_p	*repABC*-based $oriV_{pSMED01}$ was PCR amplified with primers 151+152 from *S. medicae* WSM419 cell lysate. The PCR product was 5‘ phosphorylated and inserted into SmaI-opened pK18mob2.
pSMED02_p	*repABC*-based $oriV_{pSMED02}$ was PCR amplified with primers 153+154 from *S. medicae* WSM419 cell lysate. The PCR product was 5‘ phosphorylated and forwardly inserted into SmaI-opened pK18mob2.
pNGR234b_p	*repABC*-based $oriV_{pNGR234b}$ was PCR amplified with primers 149+150 from *S. fredii* NGR234 cell lysate. The PCR product was 5‘ phosphorylated and reversely inserted into SmaI-opened pK18mob2.
pTi_p	*repABC*-based $oriV_{pTi}$ was PCR amplified with primers 137+138 from *A. tumefaciens* C58 cell lysate. The PCR product was 5‘ phosphorylated and forwardly inserted into SmaI-opened pK18mob2.

Supplementary Figures

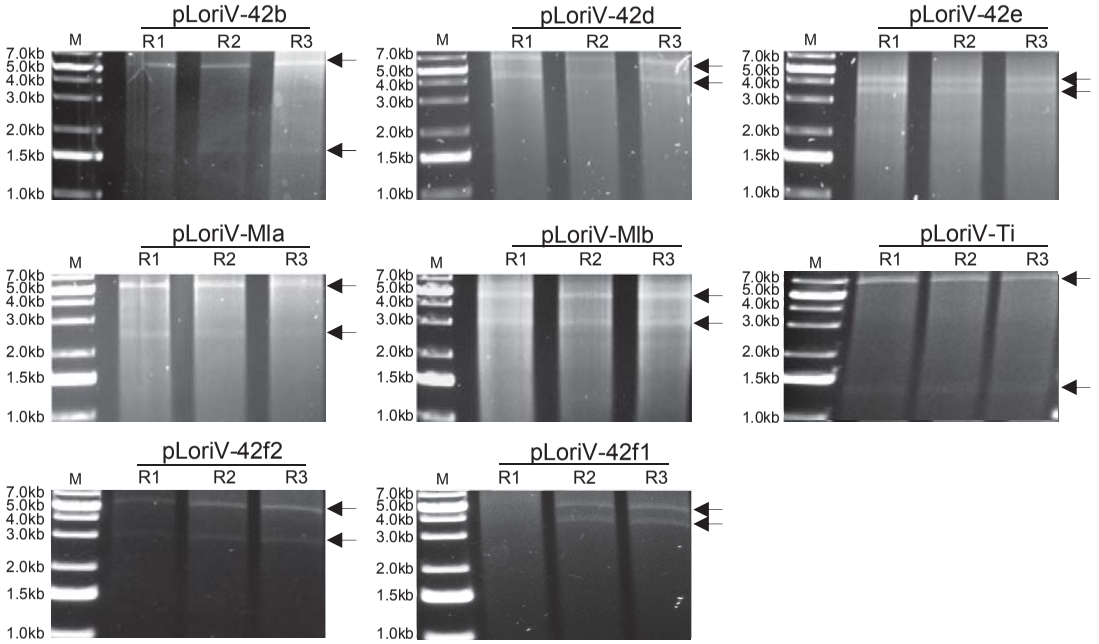


Figure S1. Electrophoresis banding patterns of oriVSm library plasmids. Plasmids were purified from three *S. meliloti* transconjugants (R1-3). Arrows indicate expected fragment sizes: pLoriV-42b (KpnI/XbaI): ~6.0+1.65kb. pLoriV-42d (SmaI/NheI): ~6.7+4.3kb. pLoriV-42e (KpnI/XhoI): ~4.2+3.5kb. pLoriV-Mla (NdeI/SmaI): ~5.5+2.5kb. pLoriV-Mlb (EcoRI/NdeI): ~4.8+3kb. pLoriV-Ti (KpnI/SmaI): ~6.4k+1.3kb. pLoriV-42f2 (KpnI/NheI): ~5.2+2.8kb. pLoriV-f42f1 (BamHI/NheI): ~4.5+3.4kb (R1 is negative). M: „GeneRuler 1 kb Plus DNA Ladder“ (Thermo Scientific).

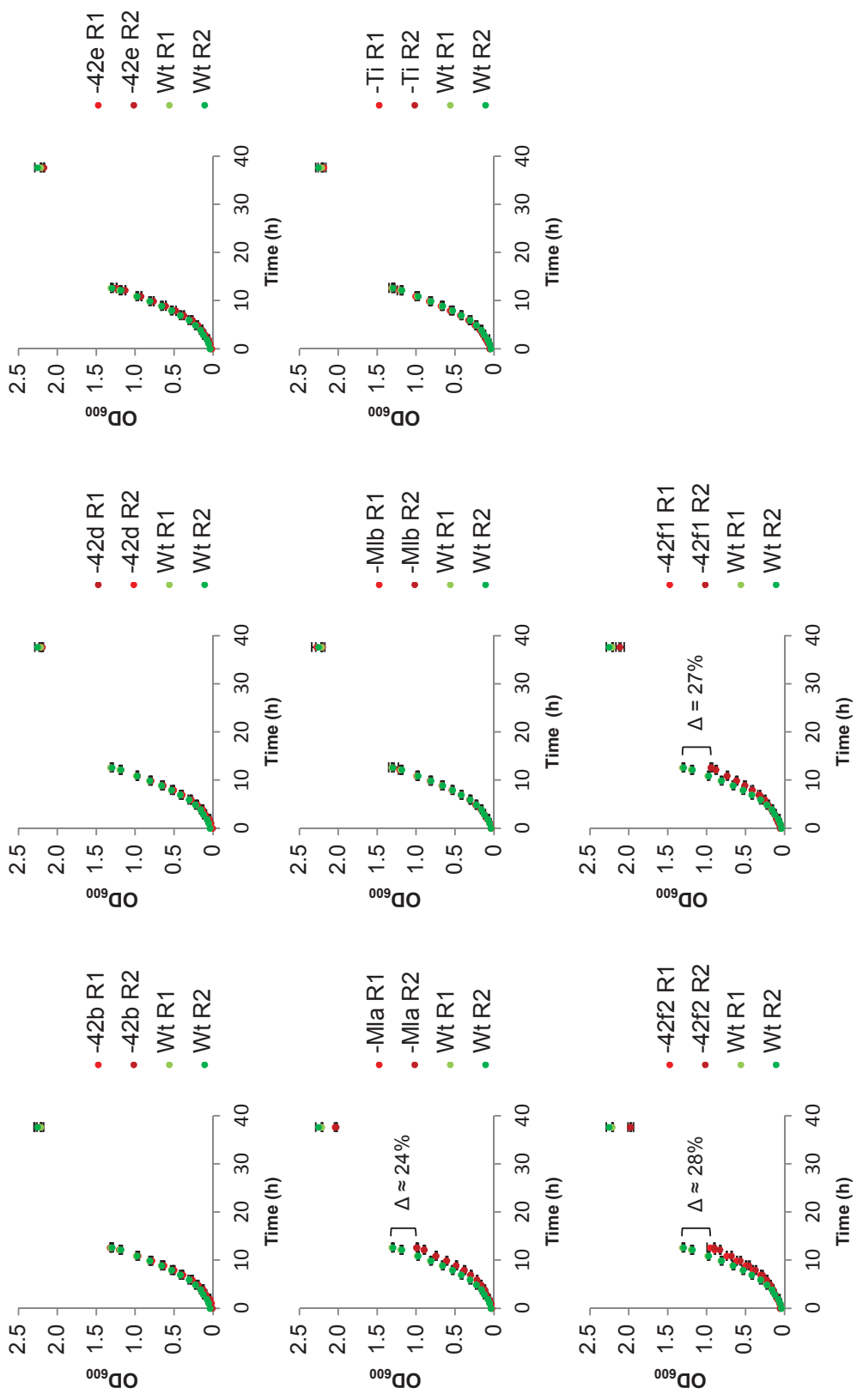


Figure S2. Growth of *S. meliloti* carrying stably propagating pLoriVSm library plasmids (pLoriV-42b, -42d, -42e -42f1, -42f2, -Mla, -Mlb and -Ti). Cultures were adjusted to an initial OD_{600} of 0.1 and incubated in TY medium supplemented with kanamycin at 30°C. Wt: *S. meliloti* MWSm8 bearing pK18mob2 genomically integrated. Experiments are based on two biological replicates (R1/2). Error bars represent the standard deviation calculated from four technical replicates.

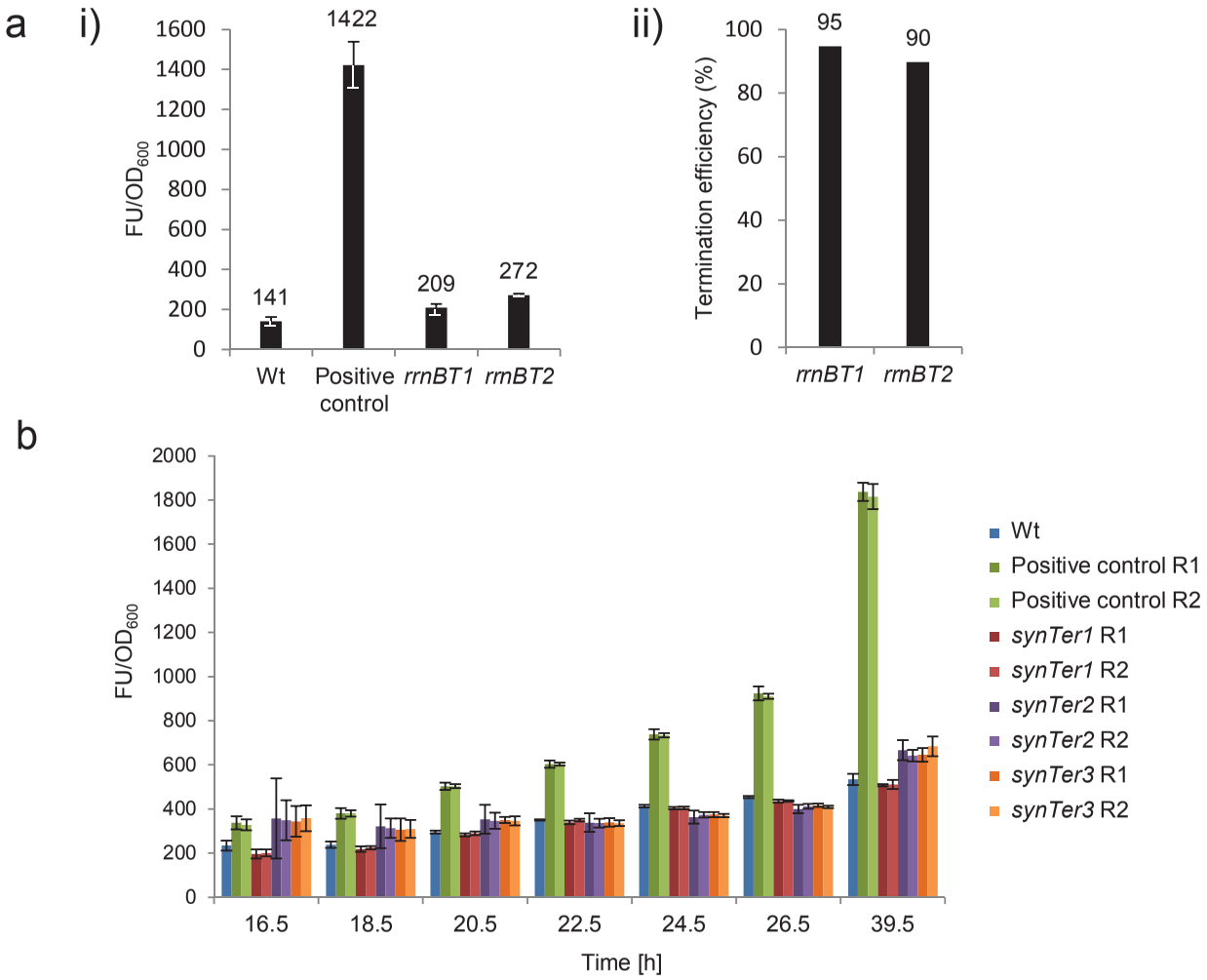


Figure S3. Activity of *E. coli* transcription terminators derived from *rrnBT1* and *rrnBT2* in *S. meliloti*. Termination efficiency was correlated with eGFP fluorescence emanating from genomically integrated *eGFP* at position Smb1167638. The cassettes comprise *eGFP* under the control of the gentamicin resistance cassette promoter P*aacC1* (Positive control) and was interrupted by *rrnBT1* (JDSm16), *rrnBT2* (JDSm17), *synTer1* (MWSm40), *synTer2* (MWSm41) and *synTer3* (MWSm42), respectively. Error bars represent the standard deviation. FU: fluorescence units. (a) (i) Termination efficiency of *rrnBT1* and *rrnBT2*. *S. meliloti* strains JDSm15-17 were incubated in TY medium supplemented with streptomycin over night at 30°C. Measurements are based on six replicates. Wt: *S. meliloti* Rm1021; Positive control: *S. meliloti* JDSm15. (ii) Termination efficiency was calculated from (i). (b) Termination efficiency of *synTer* regions 1-3. *S. meliloti* strains MWSm40-42 were incubated in TY medium supplemented with kanamycin over 39.5 h at 30°C. Measurements represent five technical replicates. R1/2: biological replicates; Wt: *S. meliloti* MWSm9; Positive control: *S. meliloti* MWSm39.

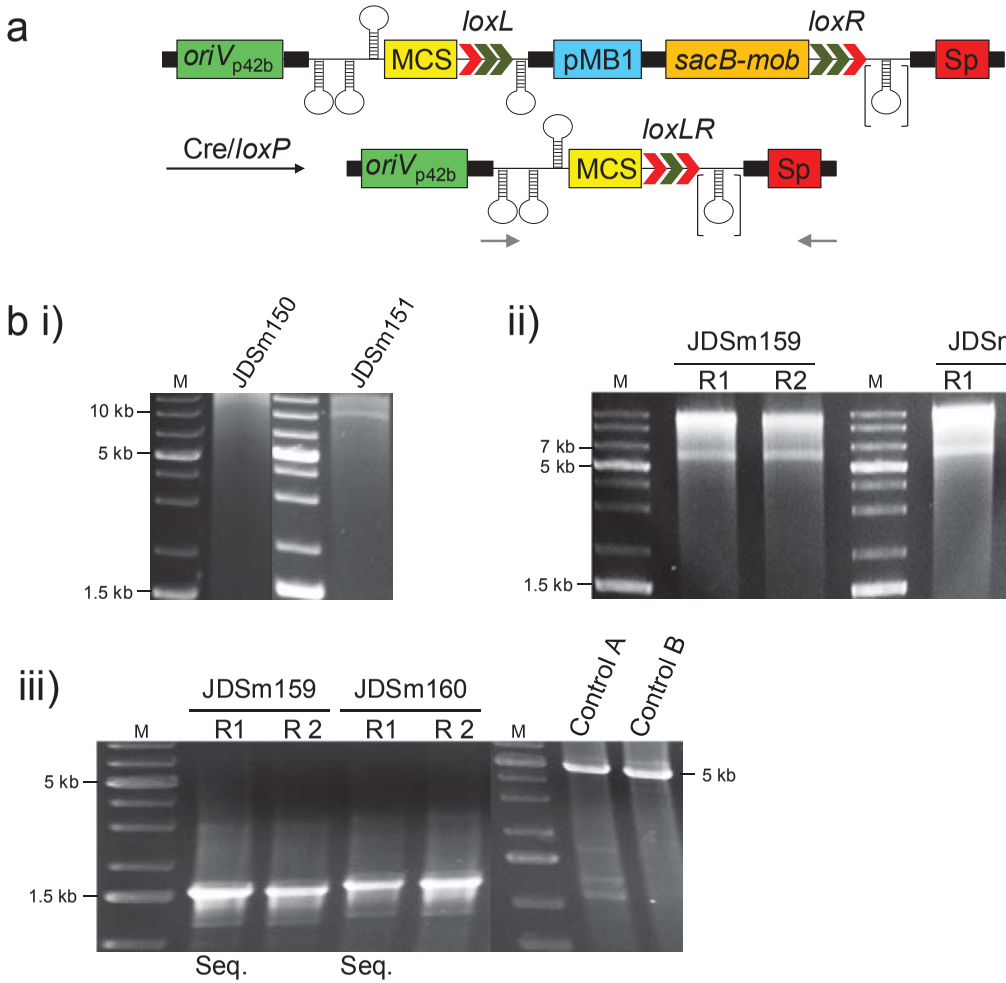


Figure S4. Cre-mediated deletion of pABC modules in *S. meliloti* strain SmCreΔhsdR. (a) Cre/*lox* mediated deletion of modules oriVEc and oriT from pABC1-loxA (design type C1) and pABC1-loxB (design type C2, containing a transcription terminator downstream of *loxR*). Successful Cre-mediated deletion and conversion of *loxL* and *loxR* into *loxLR* was verified via sequencing of PCR products (primers 496 and 503 are depicted as grew arrows). Sp: Spectinomycin resistance cassette (corresponding to module AR). (b) Examination of corresponding strains JDSm150 (carrying pABC1-loxA) and JDSm151 (carrying pABC1-loxB). (i) Plasmid DNA purified from *S. meliloti* JDSm150 (pABC1-loxA) and JDSm151 (pABC1-loxB) was ScaI digested. Electrophoresis banding pattern shows expected fragment sizes of 9.5 kb (JDSm150) and 9.6 kb (JDSm151). (ii) Fragment sizes after Cre/*loxP* recombination in JDSm150 and JDSm151, giving rise to JDSm159 (6.1 kb) and JDSm160 (6.2 kb), respectively. (iii) The ~1.5 kb deletion site from pABCs derived from JDSm159 and JDSm160 was PCR amplified with primers 496 and 503. The PCR product of R1 was sequenced with primer 496, verifying successful deletion and conversion of *loxL* and *loxR* into *loxLR*. Control A/B: Control PCR on plasmid DNA of pABC1-loxA/B. R1/2: Replicates; M: „GeneRuler 1 kb DNA Ladder“ (Thermo Scientific).

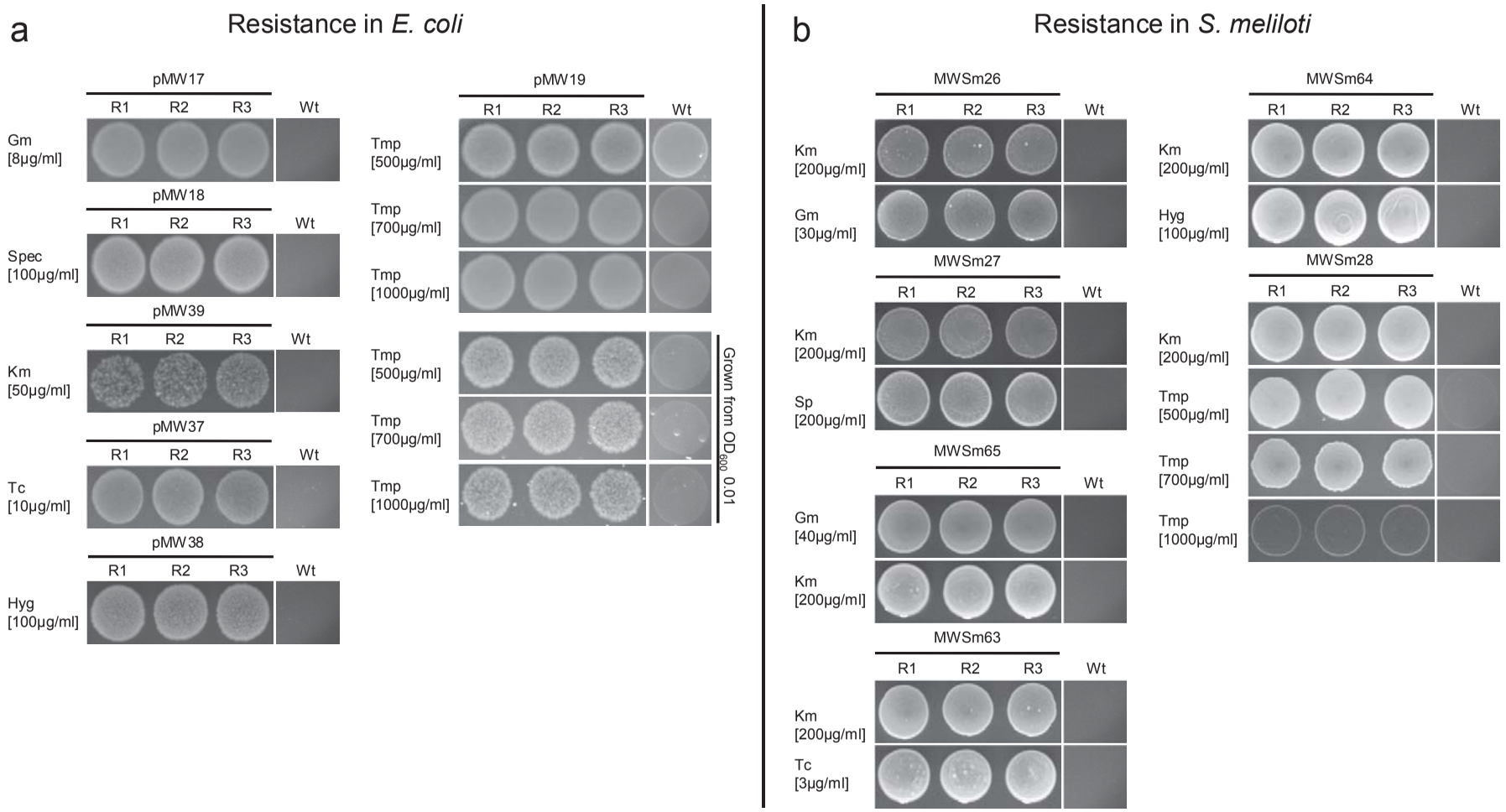


Figure S5. Examination of standardized antibiotic resistance cassettes. Constructs confer resistance to gentamicin (Gm), spectinomycin (Sp), kanamycin (Km), tetracyclin (Tc), hygromycin (Hyg) or trimethoprim (Tmp). R1-3: tested replicates. (a) *E. coli* strains carrying plasmids pMW17 (P*min2-aacC1*), pMW18 (P*min2-aadA1*), pMW39 (P*min2-nptII*), pMW37 (P*min2-tetA*), pMW38 (P*min2-hph*) and pMW19 (P*min2-tmpR*) were adjusted to an OD_{600} of 0.1 and dropped on LB agar containing the indicated antibiotic concentrations. Incubation took place overnight at 37 °C. It must be noted that proper selection for Tmp resistance was problematic when high cell densities were plated (background growth). A novel Tmp resistance gene coding for dihydrofolate reductase Tm8-3 could confer more reliable resistance to this antibiotic (Torees-Cortés et al., 2011). Wt: *E.coli* DH5α. (b) Resistance tests of *S. meliloti* strains MWSm26, MWSm27, MWSm65, MWSm63, MWSm64 and MWSm28 bearing the respective constructs genomically integrated at position Smb1166188. Cultures were adjusted to an OD_{600} of 0.1 and dropped on TY agar containing Km or Gm (dependend on the resistance conferred by the plasmid backbone) and the antibiotic to be tested, and incubated for 48 h at 30 °C. Wt: *S. meliloti* Rm1021.

Torres-Cortés, G., Millán, V., Ramírez-Saad, H. C., Nisa-Martínez, R., Toro, N., and Martínez-Abarca, F. (2011) Characterization of novel antibiotic resistance genes identified by functional metagenomics on soil samples. *Environ. Microbiol. 13*, 1101–1114.

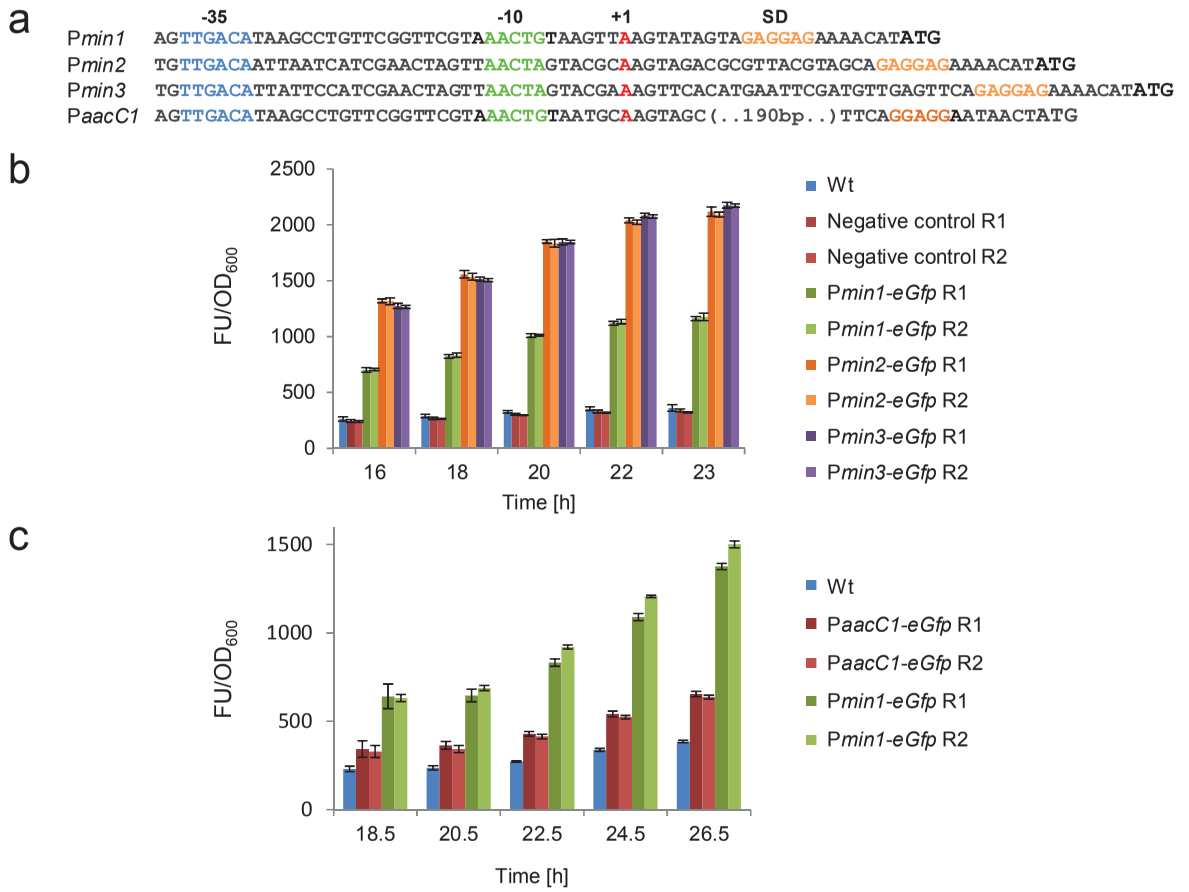


Figure S6. Examination of promoter activity of P*trp* derivatives in *S. meliloti*. (a) Sequences of synthetic promoters P*min1* to *3* and P*aacC1* derived from a gentamicin resistance cassette. Whereas P*min1* is derived from P*aacC1*, P*min2* and P*min3* correspond to P*trp* derived from *Salmonella typhimurium* (Yanofsky et al.,1981) and *Salmonella enterica* (Harrison et al., 2011), respectively. Promoter regions mainly differ with respect to the lenght of the 5‘ UTR. -35/-10: -35/-10 promoter elements; +1: transcription start site; SD: Shine-Dalgarno sequence. (b) eGFP fluorescence of *S. meliloti* strains MWSm9 (Wt), MWSm43 (negative control, promoterless *SD-eGFP* fragment), MWSm36 (P*min1-eGfp*), MWSm37 (P*min2-eGfp*) and MWSm38 (P*min3-eGfp*) which were incubated in TY medium supplemented with kanamycin at 30 °C. Plasmids are genomically integrated at position Smb1166188. Error bars represent the standard deviation calculated from five technical replicates. R1/2: biological replicates. FU: fluorescence units. (c) eGFP fluorescence of *S. meliloti* strains MWSm9 (Wt), MWSm39 (P*aacC1-eGfp*) and MWSm36 (P*min1-eGfp*) incubated in TY medium supplemented with kanamycin at 30 °C. Error bars represent the standard deviation calculated from five technical replicates.

Yanofsky, C., Platt, T., Crawford, I. P., Nichols, B. P., Christie, G. E., Horowitz, H., VanCleemput, M., and Wu, A. M. (1981) The complete nucleotide sequence of the tryptophan operon of *Escherichia coli*. *Nucleic Acids Res. 9*, 6647–6668.

Harrison, C. L., Crook, M. B., Peco, G., Long, S. R., and Griffitts, J. S. (2011) Employing site-specific recombination for conditional genetic analysis in *Sinorhizobium meliloti*. *Appl. Environ. Microbiol. 77*, 3916–3922.

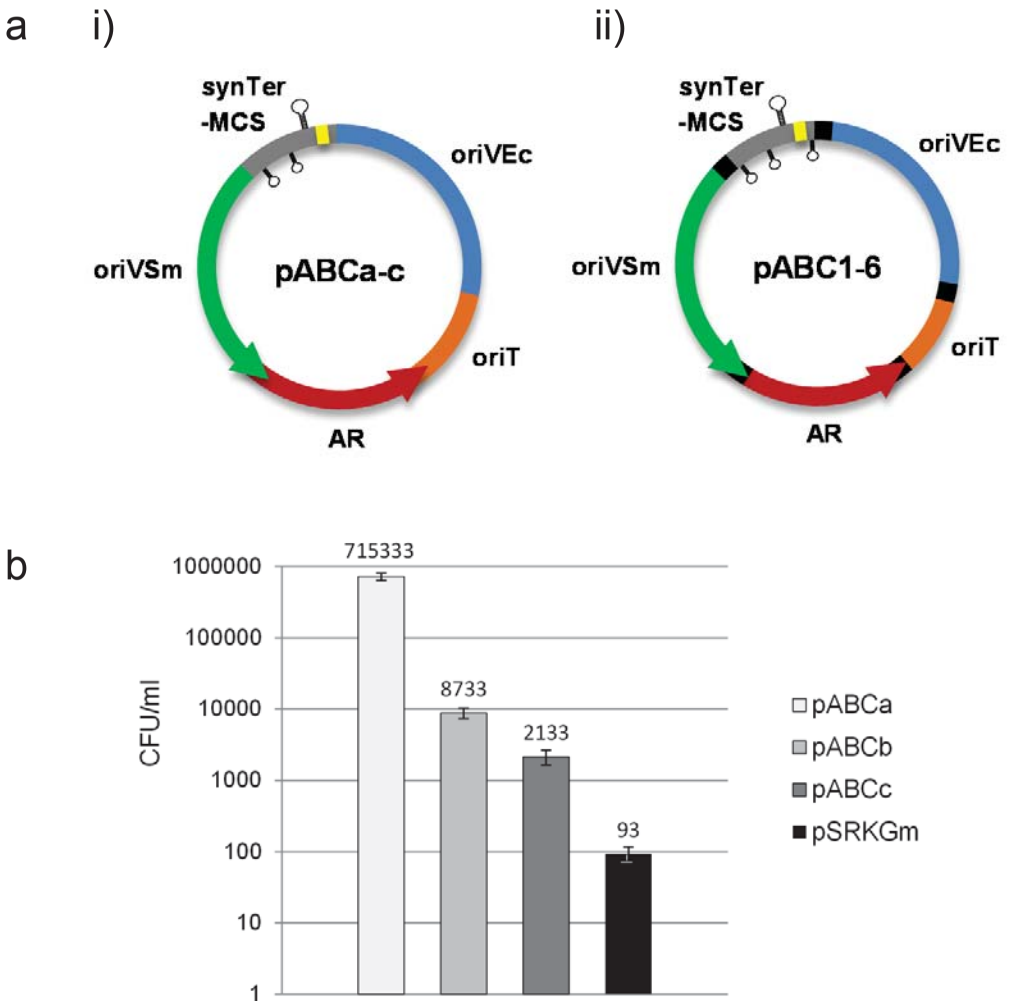


Figure S7. pABC setup and electroporation efficiency. (a) Setup of pABCs. i) pABCa-c are composed of module parts lacking linker sequences and a transcription terminator (loop structure) downstream of the MCS (yellow blox). ii) pABCs based on module library plasmids (pABC1-6) were assembled using standardized module-specific linker sequences (black boxes). (b) Electroporation efficiency of pABCa-c and pSRKGm. Plasmids were purified from *E. coli* DH5α and 50 ng DNA were used for electroporation of *S. meliloti* Rm1021 cells. Error bars represent the standard deviation calculated from three replicates. CFU: Colony forming units.

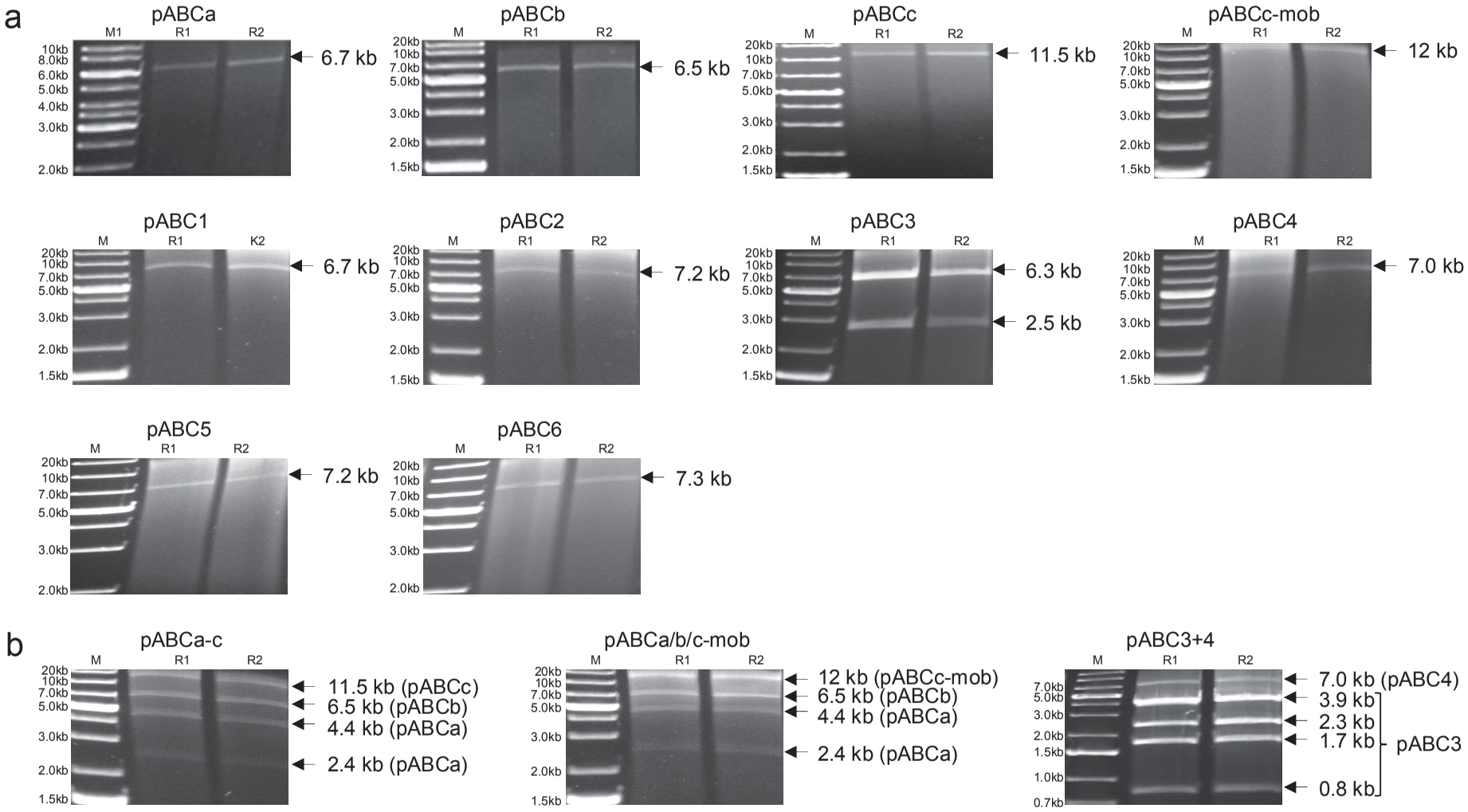
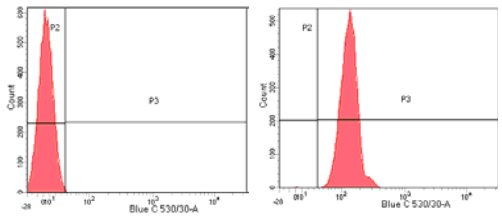


Figure S8. Electrophoresis banding pattern of restriction enzyme-digested pABCs purified from *S. meliloti*. Expected fragment sizes are indicated by arrows. (a) Digestion of single pABCs with ScaI. (b) Digestion of pABCs purified from *S. meliloti* strains harboring pABCa-c (digested with ScaI and NheI), pABCa/b/c-mob (digested with ScaI and NheI) and pABC3+4 (digested with ScaI and SmaI). M1: „GeneRuler 1 kb DNA Ladder“ (Thermo Scientific), M: „GeneRuler 1 kb Plus DNA Ladder“ (Thermo Scientific), R1/2: biological replicates.

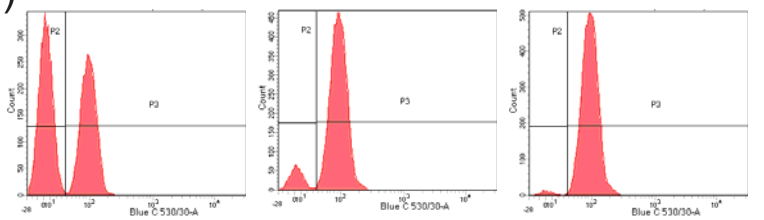
a



S. meliloti Rm1021	
Negative	Positive
99.9 %	**0.1 %**

S. meliloti MWSm37	
Negative	Positive
0.6 %	**99.4 %**

b i)

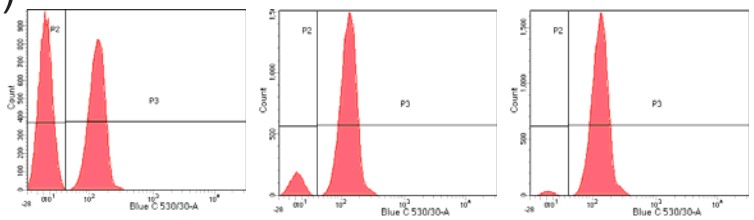


50 : 50	
Negative	Positive
52.3 %	**47.7 %**

10 : 90	
Negative	Positive
10.6 %	**89.4 %**

2 : 98	
Negative	Positive
2.7 %	**97.3 %**

ii)



50 : 50	
Negative	Positive
49.4 %	**50.6 %**

10 : 90	
Negative	Positive
10.1 %	**89.9 %**

2 : 98	
Negative	Positive
2.3 %	**97.7%**

Figure S9. Verification of Flow Cytometry measurements. *eGFP* expression is based on a P*min2-eGFP* cassette which was also used for pABC propagation stability assays. Gates were adjusted in order to distinguish between „negative“ (*S. meliloti* Rm1021; gate P2) and „postitive“ cells (*S. meliloti* MWSm37 carrying an integrated eGFP cassette; gate P3). (a) Fresh cell cultures were adjusted to an OD_{600} of 0.1 in 0.9 % NaCl for measurements. *S. meliloti* Rm1021 and MWSm37 were properly assigned to respective gates. (b) Sufficient sensitivity of measurements for the relevant cell density range was demonstrated by adjusting the ratios (negative : positive) of the respective strains. Cultures were grown to an OD_{600} of ~0.5 (i) or ~1.0 (ii), and adjusted to an OD_{600} of 0.1 in 0.9 % NaCl for measurements.

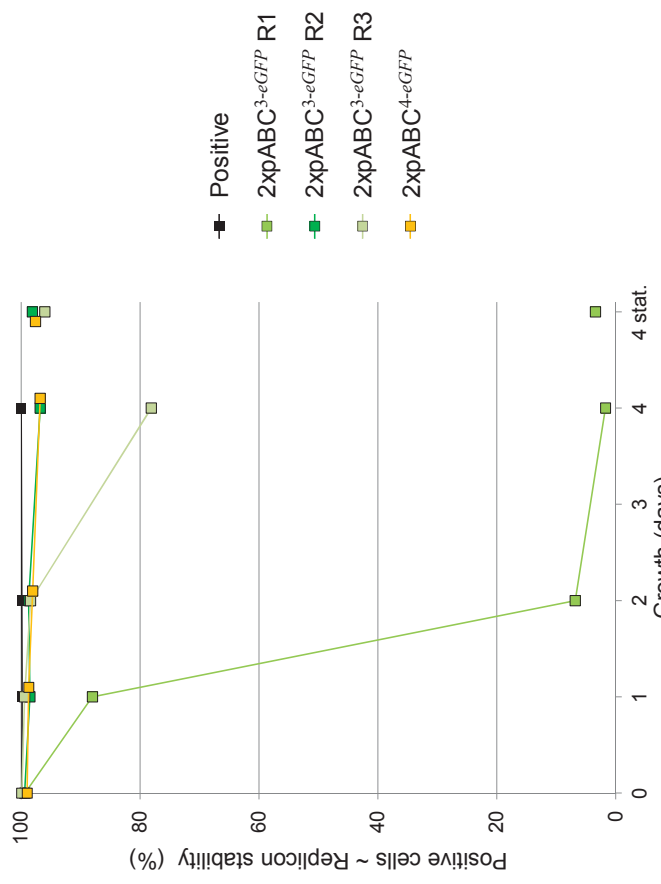


Figure S10. Propagation stability of pABC3-eGFP and pABC4-eGFP. 2xpABC$^{3\text{-}eGFP}$: *S. meliloti* Rm1021 carrying pABC3-eGFP and pABC4. 2xpABC$^{4\text{-}eGFP}$: *S. meliloti* Rm1021 carrying pABC3 and pABC4-eGFP. Measurements are based on three biological replicates. Standard deviation is below 1. Since replicates representing 2xpABC$^{3\text{-}eGFP}$ were heterogeneous, single replicates are shown (R1-3). Positive: *S. meliloti* MWSm37 (bearing a genomically integrated *eGFP* cassette).

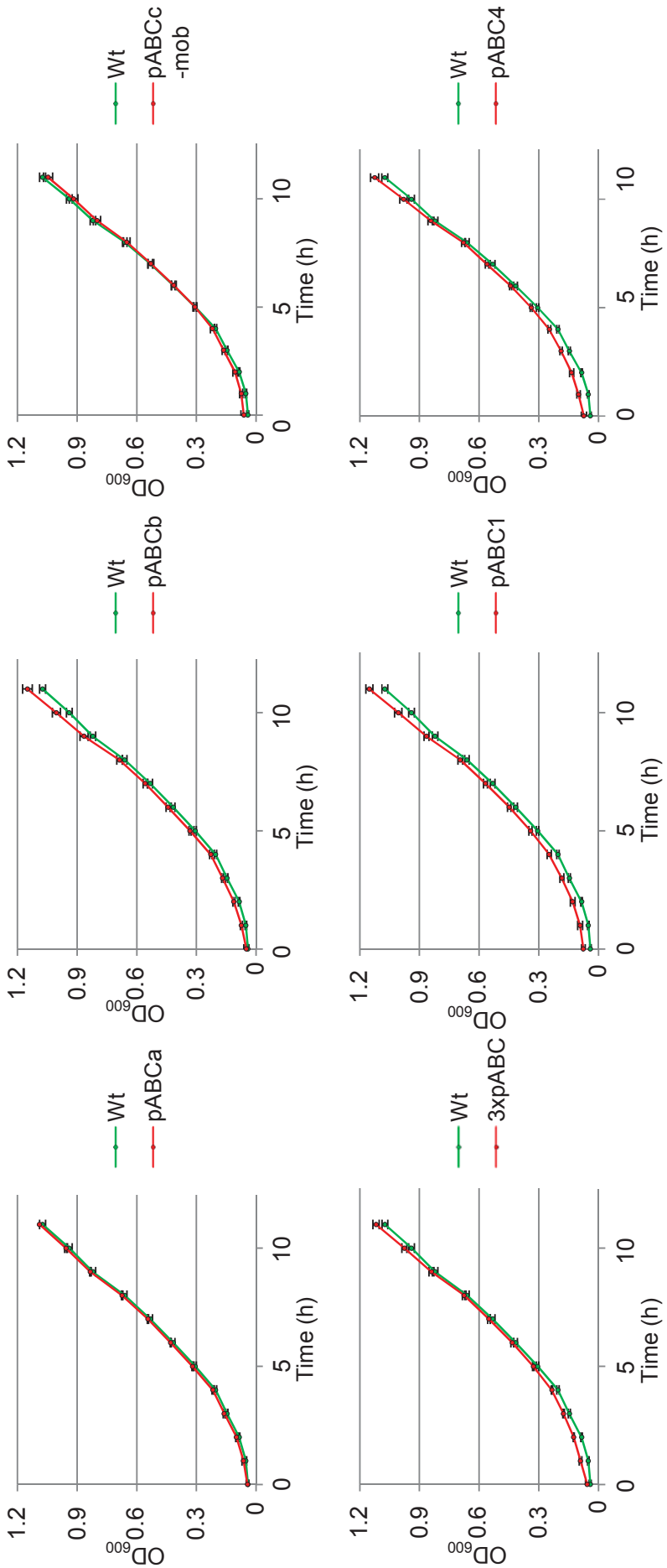
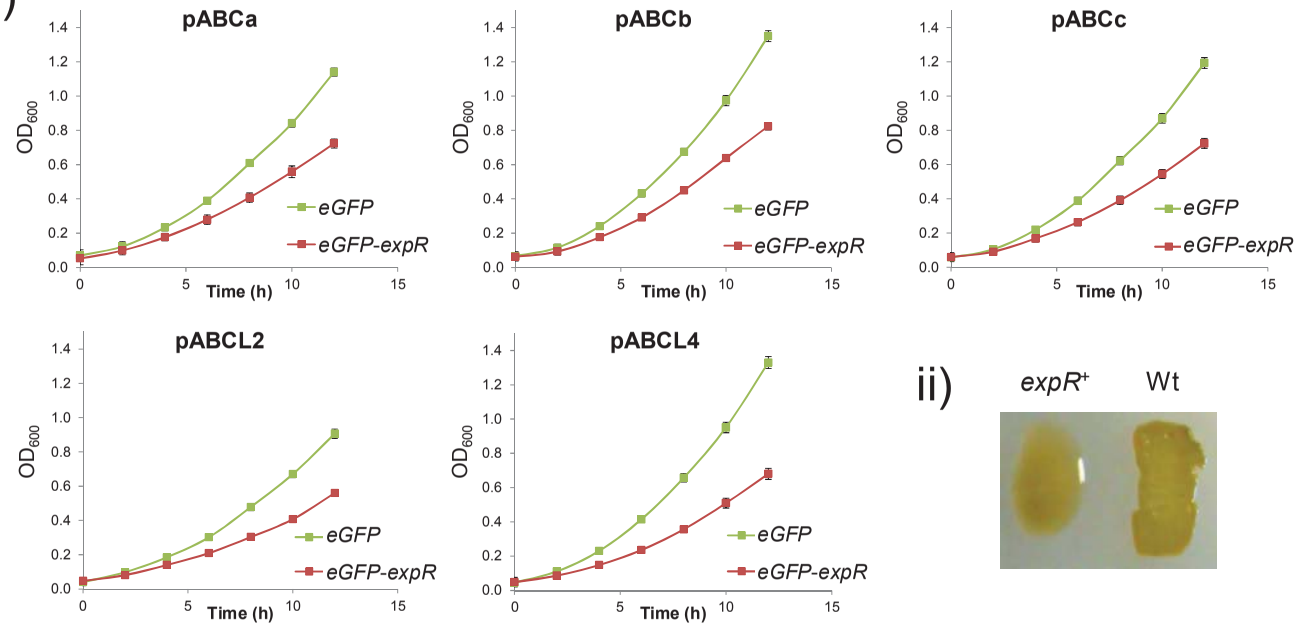


Figure S11. Growth of *S. meliloti* Rm1021 harboring stably propagating pABCs. Strains were incubated in TY medium supplemented with streptomycin and OD_{600} was measured every hour until an OD_{600} of ~1 was reached. Error bars represent the standard deviation calculated from three replicates. 3xpABC: *S. meliloti* Rm1021 carrying pABCa, pABCb and pABCc-mob. Wt: *S. meliloti* Rm1021.

a i)



ii)

b

	Mucoid colonies (%)		
***S. meliloti* carrying**	**0 h**	**48 h**	**96 h**
pABCa-eGFP-expR	100	100	100
pABCb-eGFP-expR	100	100	100
pABCc-eGFP-expR	100	100	100
pABC2-eGFP-expR	100	100	100
pABC4-eGFP-expR	100	100	91

Figure S12. Characterization of pABC-eGFP-expR derivatives. (a) Constitutive *expR* expression from pABCs is a burden for *S. meliloti*. (i) Growth of *S. meliloti* Rm1021 (Rm1021 is an *expR*$^-$ strain) carrying related derivatives of pABC-eGFP (green label) and pABC-eGFP-expR (red label). Cultures were incubated in TY medium supplemented with suitable antibiotics at 30°C and OD_{600} was measured every two hours. Error bars represent the standard deviation calculated from six replicates. (ii) *S. meliloti* Rm1021 carrying pABCa-expR (*expR*$^+$) exhibits a mucoid colony surface. Wt: *S. meliloti* Rm1021. (b) Examination of the mucoid phenotype of *S. meliloti* clones carrying pABC-eGFP-expR derivatives. Samples were isolated 0, 48 and 96 h post-inoculation during the corresponding replicon stability assay (Figure 5C). Dilutions of cell suspensions were plated on selective TY agar and emerging colonies were controlled for a mucoid phenotype (n= 10 for 0 h and 48 h; n= 11 for 96 h). A mucoid phenotype is supposed to correlate with a functional *expR* gene on the pABC. Sequencing of plasmid DNA of non-mucoid clones carrying pABC4-eGPF-expR revealed inactivation of *expR* by deletion of a 11 bp fragment at nucleotide position 172 of the *expR* coding region.

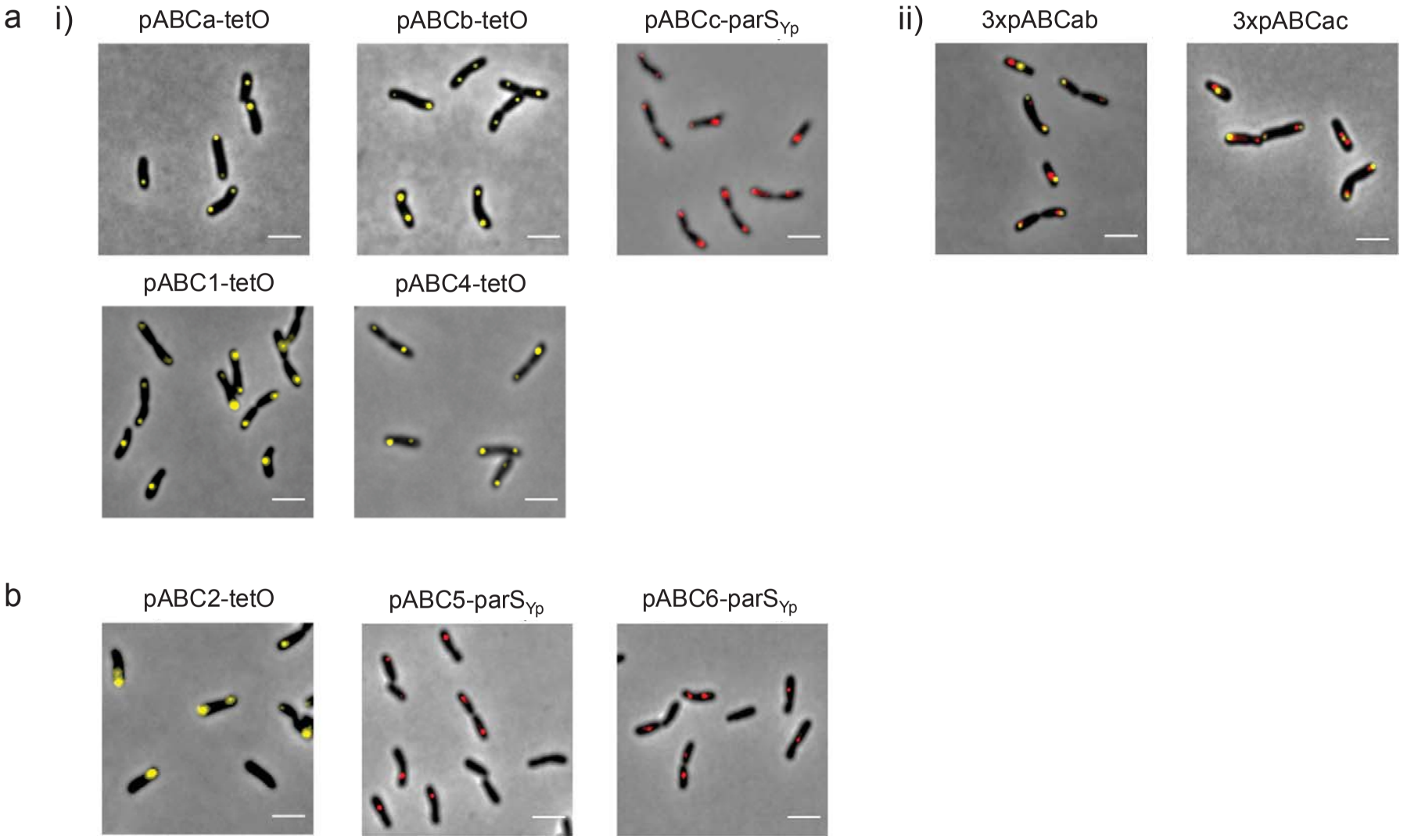


Figure S13. Fluorescence microscopy analysis of fluorescently labeled pABCs. Strains are based on *S. meliloti* JDSm106 bearing reporter genes *tetR-YFP* (yellow) and *mChr-parB*$_{Yp}$ (red) under control of the taurine-inducible promoter P*tauAB*. Prior to microscopic analysis, cultures were induced with 15 mM taurine for 4 h. Scale bar: 2.5 µm. (a) *S. meliloti* JDSm106 carrying stably propagating pABC derivatives. i) *S. meliloti* bearing a single pABC. ii) *S. meliloti* exhibiting triple pABC setups with two labeled derivatives. 3xpABCab: pABCa-tetO, pABCb-parS$_{Yp}$, pABCc; 3xpABCac: pABCa-tetO, pABCb, pABCc-parS$_{Yp}$. (b) *S. meliloti* JDSm106 carrying instable pABC derivatives.

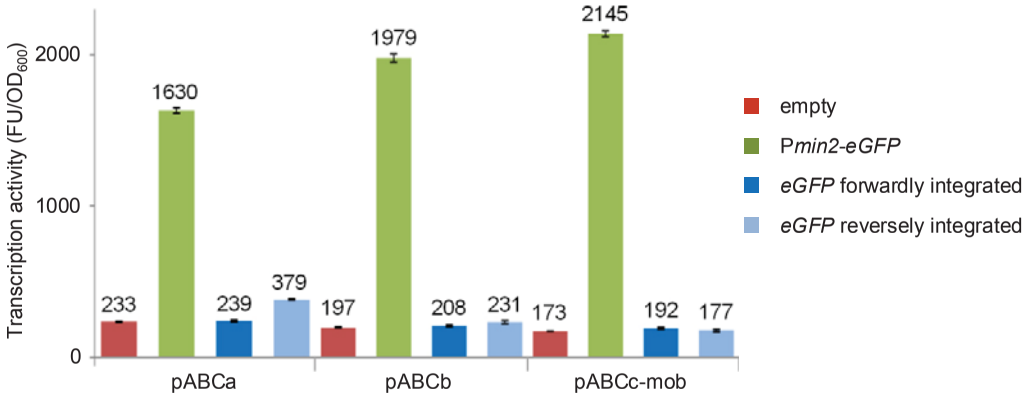
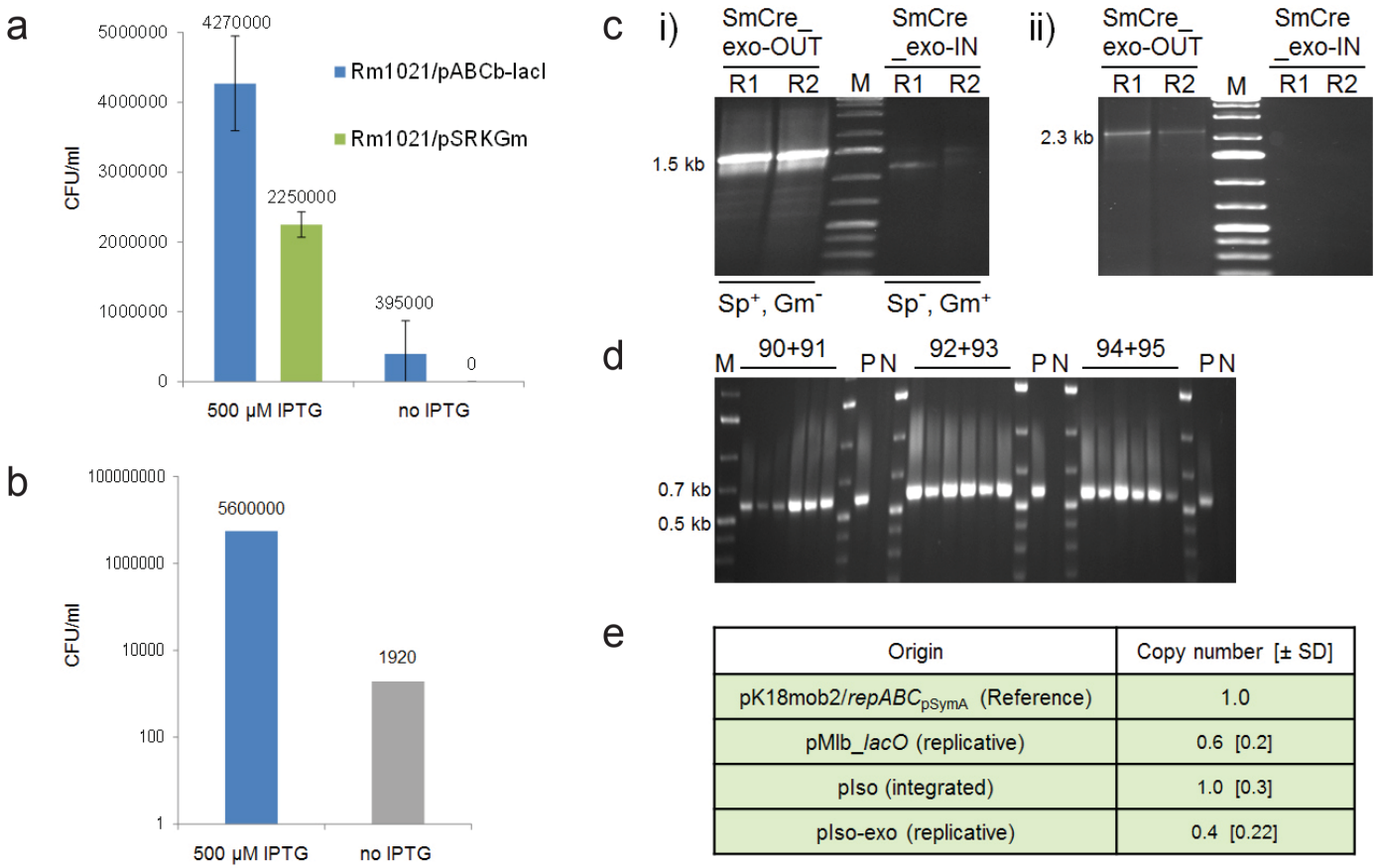


Figure S14. Analysis of transcriptional insulation of the multiple cloning site of pABCa to c. Fluorescence derived from the *eGFP* coding region inserted, either in forward or reverse orientation, into the MCS of pABCa, pABCb, and pABCc-mob was measured. *S. meliloti* Rm1021 harboring an pABC lacking the reporter gene insert in the MCS (red columns/ empty), pABCa/b/c-eGFP (green columns/ P*min2-eGFP*), pABCa/b/c-eGFPf (dark blue columns/ *eGFP* forwardly integrated) or pABCa/b/c-eGFPr (light blue colums/ *eGFP* reversely integrated) was grown in selective TY medium at 30 °C for 25 h. Error bars represent the standard deviation of EGFP fluorescence measurements calculated from six technical replicates. FU: fluorescence units.



Origin	Copy number [± SD]
pK18mob2/*repABC*$_{pSymA}$ (Reference)	1.0
pMlb_*lacO* (replicative)	0.6 [0.2]
pIso (integrated)	1.0 [0.3]
pIso-exo (replicative)	0.4 [0.22]

Figure S15. Examination of ABC cloning procedures. (a) Effect of IPTG supplementation and *lacI* expression level on conjugation efficiency of pMlb_lacO carrying the *repABC*$_{pMlb}$-*lacO* fragment. *E. coli* S17-1/pMlb_lacO was conjugated with *S. meliloti* carrying pABCb-*lacI* and pSRKGm, respectively. Error bars represent the standard deviation calculated from four replicates. (b) Conjugation efficiency of pISO into *S. meliloti* JDSm201 (precursor of SmCre_exo-IN). Dilutions of mating mixtures were plated on selective TY agar either supplemented with 500 or 0 µM IPTG. Values are based on duplicates. (c) Colony PCR on *S. meliloti* strains SmCre_exo-OUT and SmCre_exo-IN. R1/2: independent replicates. i) PCR amplification of the 1460 bp deletion site on pSymB with primers 94 and Rev. Sequencing of PCR products verified conversion of *loxL* into *loxLR*. As expected, Cre recombination caused a resistance switch from gentamicin (Gm$^+$, SmCre_exo-IN) to spectinomycin resistance (Sp$^+$, SmCre_exo-OUT). ii) PCR amplification of the 2233 bp fusion site on pIso-exo with primers 680 and 128. Sequencing of PCR products verified conversion of *loxR* into *loxP*. M: „GeneRuler 1 kb DNA Ladder“ (Thermo Scientific). (d) qPCR-based copy number determination of *repABC*$_{pMlb}$-based constructs using cell lysates of *S. meliloti* Rm1021 carrying pMlb_lacO (replicative), *S. meliloti* SmCre_exo-IN (pIso integrated) and *S. meliloti* SmCre_exo-OUT (pIso-exo replicative). *S. meliloti* Sma818 carrying pK18mob2/*repABC*$_{pSymA}$ served as reference strain. For qPCR, primer combinations 78+79 (binding the chromosome) and 252+253 (binding the *mob* site) were applied. Values < 1 indicate instable propagation since examined cell suspensions are supposed to contain progeny lacking the analyzed plasmid. SD: Standard deviation. (e) PCR amplification of ~550 bp fragments from transconjugants of *A. tumefaciens*/pIso-exo with primer combinations specific for *exoB* (90+91), *exoI* (92+93) and *exsA* (94+95). P: pIso-exo purified from *E. coli*; N: *A. tumefaciens* C58 exoB. M: „GeneRuler 1 kb DNA Ladder“ (Thermo Scientific).

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertation mit dem Titel „Etablierung des stickstofffixierenden Alphaproteobakteriums *Sinorhizobium meliloti* als Chassis-Organismus in der Synthetischen Mikrobiologie“ selbstständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinem sonstigen Prüfungszwecken gedient.

_____ _____

(Ort / Datum) (Johannes Döhlemann)

Danksagung

Abschließend möchte ich allen Menschen danken, die einen Anteil an der Entstehung dieser Dissertation hatten.

Mein besonderer Dank gilt Frau Prof. Dr. Anke Becker für die Bereitstellung dieses spannenden Forschungsthemas, die Betreuung meiner Dissertation und den anregenden, wissenschaftlichen Diskurs.

Weiterhin möchte ich Herrn Prof. Dr. Martin Thanbichler für die Übernahme des Zweitgutachtens danken.

Ein großer Dank gilt außerdem den aktuellen und ehemaligen Mitgliedern der Arbeitsgruppen Becker und Waldminghaus. Ohne das freundliche Klima und die kooperative Zusammenarbeit wäre diese Arbeit sicherlich weniger angenehm verlaufen. Ich möchte mich bei Liza für ihre geduldige Beratung in praktischen und fachlichen Fragen bedanken. Für die wertvollen Ratschläge bei genetischen Anliegen möchte ich zudem Matthew danken. Ein großer Dank gebührt auch Benny und Jan für die Beantwortung der großen und kleinen Fragen am Mikroskop. Für die fächmännische Instandsetzung der Pulsfeldgelelektrophorese bedanke ich mich zudem bei Patrick. Einen besonderen Dank möchte ich meinen Masterstudenten Marcel, Meike und Carina aussprechen. Danke für eure Motivation, den Fleiß und Spaß, mit dem ihr diese Arbeit entscheidend unterstützt habt.

Ein großer Dank geht an Julius, Till, Ralf und Daniel für das notwendige Maß Ablenkung jenseits der Wissenschaft und Marburg.

Ich möchte mich auch herzlich bei meinen Eltern und Geschwistern für die familiäre Unterstützung bedanken. Mein ganz besonderer Dank gilt Julia, die mir auch in den schwierigsten Stunden den nötigen Rückhalt gab.

Der Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt